

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 22.0606 Seconds
(without alignments)
179.063 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1 KATMQCGENEKYDSCGSKC.....VSAEDCLNDMDFIYFGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	117	24.1	77	1 CVP6 PIMHY	Q8t0w0 pimpla hypo
2	110.5	22.7	56	1 AMCI_APTIME	P56682 apis mellif
3	103	21.2	2812	1 ZAN HUMAN	Q9y493 homo sapien
4	100	20.6	5376	1 ZAN MOUSE	O88799 mus musculu
5	94.5	19.4	4560	1 LR2 RAT	P98158 rattus norv
6	92.5	19.0	63	1 ICB1_ASCSU	P07851 ascaris suu
7	92.5	19.0	4753	1 LRP_CAEEL	Q04833 caenorhabdi
8	87.5	18.0	2813	1 VWF CANFA	Q28295 canis fami
9	87	17.9	115	1 A62F_DROME	O46202 drosophila
10	86.5	17.8	4655	1 LR2 HUMAN	P98164 homo sapien
11	85.5	17.6	65	1 ICB2_ASCSU	P07852 ascaris suu
12	85.5	17.6	863	1 LDVR_CHICK	P98165 gallus gall
13	82.5	17.0	873	1 LDVR_HUMAN	P98155 homo sapien
14	82.5	17.0	873	1 LDVR_MOUSE	P98156 mus musculu
15	82.5	17.0	873	1 LDVR_RABIT	P35953 oryctolagus
16	82.5	17.0	873	1 LDVR RAT	P98166 rattus norv
17	82.5	17.0	2282	1 ZAN_RABIT	P57999 oryctolagus
18	82	16.9	62	1 ITR1_ASCSU	P13998 ascaris suu
19	81.5	16.8	1808	1 TENA_CHICK	P10039 gallus gall
20	81	16.7	62	1 MT4 CANFA	Q9ukf5 canis fami
21	80	16.5	790	1 AD30 HUMAN	Q9ukf2 homo sapien
22	79.5	16.4	2813	1 VWF HUMAN	P04275 homo sapien
23	79.5	16.4	2871	1 FBN1_HUMAN	P35555 homo sapien
24	79.5	16.4	2871	1 FBN1_PIG	Q9tv36 sus scrofa
25	78.5	16.2	77	1 ASP2_ANISI	O77417 anisakis si
26	78.5	16.2	798	1 ITR1_XENLA	P12606 xenopus lae
27	78.5	16.2	937	1 VWF BOVIN	P80012 bos taurus
28	78.5	16.2	3375	1 UN52_CAEEL	Q06561 caenorhabdi
29	78	16.0	289	1 TNR5_MOUSE	P27512 mus musculu
30	78	16.0	1107	1 YLKT_CAEEL	P41950 caenorhabdi
31	78	16.0	3133	1 HKCT_BOMMO	P98092 bombyx mori
32	77.5	15.9	474	1 VSM5_TRYBB	P26333 trypanosoma
33	77.5	15.9	2871	1 FBN1_BOVIN	P98133 bos taurus

ALIGNMENTS

RESULT 1

ID	CVP6 PIMHY	STANDARD;	PRT;	77 AA.
DT	Q8t0w0:			
DT	15-SEP-2003 (Rel. 42, Created)			
DT	15-SEP-2003 (Rel. 42, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Cysteine-rich venom protein 6 precursor.			
GN	CVP6.			
OS	Pimpla hypochondriaca (Parasitoid wasp).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Hymenoptera; Apocrita; Ichneumonidae;			
OC	Ichneumonidae; Pimplinae; Pimpla.			
OX	NCBI TaxID=135724;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Venom;			
RA	Parkinson N.M.;			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
CC	-!- SUBCELLULAR LOCATION: Secreted (Potential).			
CC	-!- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; AJ438997; CAD27742.1; -
InterPro; IPR002919; TIL_Cysrich.

Pfam; PF01826; TIL; 1.

Signal.

FT CHAIN 1 21

FT DOMAIN 22 77

FT DISULFID 24 56

FT DISULFID 33 52

FT DISULFID 37 48

FT DISULFID 41 76

FT DISULFID 58 70

SQ SEQUENCE 77 AA; 8184 MW; 6889CB5F40D63DD6 CRC64;

Query Match 24.1%; Score 117; DB 1; Length 77;

Best Local Similarity 35.4%; Pred. No. 4.1e-05;

Matches 23; Conservative 8; Mismatches 22; Indels 12; Gaps 3;

OY 6 CGENKDYDSCGSKCKKCKYDGVVEEDDEPNVPCIVRVCHQDCVCEGFRNKDKCV 65

DB 24 CGPNRVYKSCGT-CGPETC-----ENPDPDC-DRACHQGCFCGKLLQIDGNCI 71

OY 66 SAEDC 70

DB 72 SPDKC 76

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC
CC
CC
DR EMBL; AF332975; AAK01431.1; -
DR EMBL; AF332976; AAK01432.1; -
DR EMBL; AF332977; AAK01433.1; -
DR EMBL; AF332978; AAK01434.1; -
DR EMBL; AF332979; AAK01435.1; -
DR EMBL; AF332980; AAK01436.1; -
DR EMBL; AY046055; AAL04410.1; -
DR EMBL; AY046055; AAL04411.1; -
DR EMBL; AY046055; AAL04412.1; -
DR EMBL; AY046055; AAL04413.1; -
DR EMBL; AY046055; AAL04414.1; -
DR EMBL; AY046055; AAL04415.1; -
DR EMBL; AF053356; AAK78790.1; ALT_SEQ.
DR EMBL; AF312032; AAK21011.1; -
DR EMBL; U83191; AAC51208.1; -
DR Genew; HGNC:12857; ZAN.
DR MIM; 602372; -
DR GO; GO:0016324; C:apical plasma membrane; NAS.
DR GO; GO:0007339; P:binding of sperm to zona pellucida; NAS.
DR GO; GO:0008037; P:cell recognition; NAS.
DR GO; GO:0016337; P:cell-cell adhesion; NAS.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000998; NAM domain.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR003328; TILa_Cysrich.
DR InterPro; IPR001007; VWF_C.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF00629; NAM; 3.
DR Pfam; PF01826; TIL; 4.
DR Pfam; PF02345; TILa; 5.
DR Pfam; PF00094; vwd; 4.
DR SMART; SM00181; EGF; 4.
DR SMART; SM00137; NAM; 3.
DR SMART; SM00214; VMC; 4.
DR SMART; SM00215; VMC_out; 4.
DR SMART; SM00216; VMD; 4.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS00740; NAM_1; 1.
DR PROSITE; PS00060; NAM_2; 3.
KW Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
KW Repeat; Alternative splicing.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 2812 ZONADHESIN.
FT DOMAIN 18 2757 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 2758 2778 POTENTIAL.
FT DOMAIN 2779 2812 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 39 204 MAM 1.
FT DOMAIN 209 368 MAM 2.
FT DOMAIN 371 536 MAM 3.
FT DOMAIN 573 1041 66 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
(MUCIN-LIKE DOMAIN).
FT DOMAIN 1156 1308 VWF1.
FT DOMAIN 1542 1697 VWF2.
FT DOMAIN 1931 2086 VWF3.
FT DOMAIN 2331 2485 VWF4.
FT DOMAIN 2708 2744 EGF-LIKE.
FT DISULFID 2712 2723 BY SIMILARITY.
FT DISULFID 2717 2732 BY SIMILARITY.
FT DISULFID 2734 2743 BY SIMILARITY.
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1112 1112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1188 1188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1695 1695 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1804 1804 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1900 1900 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1946 1946 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2203 2203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2542 2542 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 2701 2701 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 2597 2601 HGVS -> WAAAF (in isoform 5).
FT VARSPLIC 2602 2812 /FTID=VSP 001420.
FT VARSPLIC 2597 2617 Missing (in isoform 5).
FT VARSPLIC 2618 2708 /FTID=VSP 001421.
FT VARSPLIC 2597 2624 HGVSRYHISELYDTPSLCQPRRG -> YAILCOBAGALAGWR
DRTLC (in isoform 6).
FT VARSPLIC 2618 2708 /FTID=VSP 001422.
FT VARSPLIC 2597 2624 Missing (in isoform 6).
FT VARSPLIC 2618 2708 /FTID=VSP 001423.
FT VARSPLIC 2597 2624 HGVSRYHISELYDTPSLCQPRRG -> YAILCOBAG
AALAGWRDRLCAGQQLSD (in isoform 4).
FT VARSPLIC 2625 2812 /FTID=VSP 001424.
FT VARSPLIC 2597 2636 Missing (in isoform 4).
FT VARSPLIC 2597 2636 HGVSRYHISELYDTPSLCQPRRGRLRGLRQH
-> YAILCOBAGALAGWRDRLCAMECPAGTGYSCMTFC
PASCANLADPGCEGCPVEGCAD (in isoform 7).
FT VARSPLIC 2663 2666 /FTID=VSP 001426.
FT VARSPLIC 2597 2689 LGSS -> VRAGRRPWGAEPARRARPGMELERLLALLPFL
AGQQ (in isoform 7).
FT VARSPLIC 2690 2812 /FTID=VSP 001427.
FT VARSPLIC 2597 2724 HGVSRYHISELYDTPSLCQPRRGRLRGLRQH
RLCLQWHPPEPLADCGCTSNGLYQIGSSFLTECSQRTIC
ASSRLLCEPF -> YAILCOBAGALAGWRDRLCAMECP
ACTIYQSCMTPCPASCANLADPGCEGCPVEGCASIPGYAY
SCTQSLPWLTVAAPMASTTSHAAF (in isoform
2).
FT VARSPLIC 2690 2812 /FTID=VSP 001428.
FT VARSPLIC 2597 2724 Missing (in isoform 2).
FT VARSPLIC 2690 2812 /FTID=VSP 001429.
FT VARSPLIC 2597 2724 HGVSRYHISELYDTPSLCQPRRGRLRGLRQH
RLCLQWHPPEPLADCGCTSNGLYQIGSSFLTECSQRTIC
ASSRLLCEPFAGEVCTLNGHTQGGFPESFCLQPCQN
DQCR -> YAILCOBAGALAGWRDRLCAMECPAGTGY
SCMTPCPASCANLADPGCEGCPVEGCASIPGYAYSGTSL
PWLTVAPAMASTTSHAAF (in isoform 1).
FT VARSPLIC 2725 2812 /FTID=VSP 001430.
FT CONFLICT 430 430 Missing (in isoform 1).
FT H -> Q (IN REF. 1; AAK01431/AAK01432/
AAK01433/AAK01434/AAK01435/AAK01436/AAK01437/AAK01438/AAK01439/AAK01440/AAK01441/AAK01442/AAK01443/AAK01444/AAK01445/AAK01446/AAK01447/AAK01448/AAK01449/AAK01450/AAK01451/AAK01452/AAK01453/AAK01454/AAK01455/AAK01456/AAK01457/AAK01458/AAK01459/AAK01460/AAK01461/AAK01462/AAK01463/AAK01464/AAK01465/AAK01466/AAK01467/AAK01468/AAK01469/AAK01470/AAK01471/AAK01472/AAK01473/AAK01474/AAK01475/AAK01476/AAK01477/AAK01478/AAK01479/AAK01480/AAK01481/AAK01482/AAK01483/AAK01484/AAK01485/AAK01486/AAK01487/AAK01488/AAK01489/AAK01490/AAK01491/AAK01492/AAK01493/AAK01494/AAK01495/AAK01496/AAK01497/AAK01498/AAK01499/AAK01500/AAK01501/AAK01502/AAK01503/AAK01504/AAK01505/AAK01506/AAK01507/AAK01508/AAK01509/AAK01510/AAK01511/AAK01512/AAK01513/AAK01514/AAK01515/AAK01516/AAK01517/AAK01518/AAK01519/AAK01520/AAK01521/AAK01522/AAK01523/AAK01524/AAK01525/AAK01526/AAK01527/AAK01528/AAK01529/AAK01530/AAK01531/AAK01532/AAK01533/AAK01534/AAK01535/AAK01536/AAK01537/AAK01538/AAK01539/AAK01540/AAK01541/AAK01542/AAK01543/AAK01544/AAK01545/AAK01546/AAK01547/AAK01548/AAK01549/AAK01550/AAK01551/AAK01552/AAK01553/AAK01554/AAK01555/AAK01556/AAK01557/AAK01558/AAK01559/AAK01560/AAK01561/AAK01562/AAK01563/AAK01564/AAK01565/AAK01566/AAK01567/AAK01568/AAK01569/AAK01570/AAK01571/AAK01572/AAK01573/AAK01574/AAK01575/AAK01576/AAK01577/AAK01578/AAK01579/AAK01580/AAK01581/AAK01582/AAK01583/AAK01584/AAK01585/AAK01586/AAK01587/AAK01588/AAK01589/AAK01590/AAK01591/AAK01592/AAK01593/AAK01594/AAK01595/AAK01596/AAK01597/AAK01598/AAK01599/AAK01600/AAK01601/AAK01602/AAK01603/AAK01604/AAK01605/AAK01606/AAK01607/AAK01608/AAK01609/AAK01610/AAK01611/AAK01612/AAK01613/AAK01614/AAK01615/AAK01616/AAK01617/AAK01618/AAK01619/AAK01620/AAK01621/AAK01622/AAK01623/AAK01624/AAK01625/AAK01626/AAK01627/AAK01628/AAK01629/AAK01630/AAK01631/AAK01632/AAK01633/AAK01634/AAK01635/AAK01636/AAK01637/AAK01638/AAK01639/AAK01640/AAK01641/AAK01642/AAK01643/AAK01644/AAK01645/AAK01646/AAK01647/AAK01648/AAK01649/AAK01650/AAK01651/AAK01652/AAK01653/AAK01654/AAK01655/AAK01656/AAK01657/AAK01658/AAK01659/AAK01660/AAK01661/AAK01662/AAK01663/AAK01664/AAK01665/AAK01666/AAK01667/AAK01668/AAK01669/AAK01670/AAK01671/AAK01672/AAK01673/AAK01674/AAK01675/AAK01676/AAK01677/AAK01678/AAK01679/AAK01680/AAK01681/AAK01682/AAK01683/AAK01684/AAK01685/AAK01686/AAK01687/AAK01688/AAK01689/AAK01690/AAK01691/AAK01692/AAK01693/AAK01694/AAK01695/AAK01696/AAK01697/AAK01698/AAK01699/AAK01700/AAK01701/AAK01702/AAK01703/AAK01704/AAK01705/AAK01706/AAK01707/AAK01708/AAK01709/AAK01710/AAK01711/AAK01712/AAK01713/AAK01714/AAK01715/AAK01716/AAK01717/AAK01718/AAK01719/AAK01720/AAK01721/AAK01722/AAK01723/AAK01724/AAK01725/AAK01726/AAK01727/AAK01728/AAK01729/AAK01730/AAK01731/AAK01732/AAK01733/AAK01734/AAK01735/AAK01736/AAK01737/AAK01738/AAK01739/AAK01740/AAK01741/AAK01742/AAK01743/AAK01744/AAK01745/AAK01746/AAK01747/AAK01748/AAK01749/AAK01750/AAK01751/AAK01752/AAK01753/AAK01754/AAK01755/AAK01756/AAK01757/AAK01758/AAK01759/AAK01760/AAK01761/AAK01762/AAK01763/AAK01764/AAK01765/AAK01766/AAK01767/AAK01768/AAK01769/AAK01770/AAK01771/AAK01772/AAK01773/AAK01774/AAK01775/AAK01776/AAK01777/AAK01778/AAK01779/AAK01780/AAK01781/AAK01782/AAK01783/AAK01784/AAK01785/AAK01786/AAK01787/AAK01788/AAK01789/AAK01790/AAK01791/AAK01792/AAK01793/AAK01794/AAK01795/AAK01796/AAK01797/AAK01798/AAK01799/AAK01800/AAK01801/AAK01802/AAK01803/AAK01804/AAK01805/AAK01806/AAK01807/AAK01808/AAK01809/AAK01810/AAK01811/AAK01812/AAK01813/AAK01814/AAK01815/AAK01816/AAK01817/AAK01818/AAK01819/AAK01820/AAK01821/AAK01822/AAK01823/AAK01824/AAK01825/AAK01826/AAK01827/AAK01828/AAK01829/AAK01830/AAK01831/AAK01832/AAK01833/AAK01834/AAK01835/AAK01836/AAK01837/AAK01838/AAK01839/AAK01840/AAK01841/AAK01842/AAK01843/AAK01844/AAK01845/AAK01846/AAK01847/AAK01848/AAK01849/AAK01850/AAK01851/AAK01852/AAK01853/AAK01854/AAK01855/AAK01856/AAK01857/AAK01858/AAK01859/AAK01860/AAK01861/AAK01862/AAK01863/AAK01864/AAK01865/AAK01866/AAK01867/AAK01868/AAK01869/AAK01870/AAK01871/AAK01872/AAK01873/AAK01874/AAK01875/AAK01876/AAK01877/AAK01878/AAK01879/AAK01880/AAK01881/AAK01882/AAK01883/AAK01884/AAK01885/AAK01886/AAK01887/AAK01888/AAK01889/AAK01890/AAK01891/AAK01892/AAK01893/AAK01894/AAK01895/AAK01896/AAK01897/AAK01898/AAK01899/AAK01900/AAK01901/AAK01902/AAK01903/AAK01904/AAK01905/AAK01906/AAK01907/AAK01908/AAK01909/AAK01910/AAK01911/AAK01912/AAK01913/AAK01914/AAK01915/AAK01916/AAK01917/AAK01918/AAK01919/AAK01920/AAK01921/AAK01922/AAK01923/AAK01924/AAK01925/AAK01926/AAK01927/AAK01928/AAK01929/AAK01930/AAK01931/AAK01932/AAK01933/AAK01934/AAK01935/AAK01936/AAK01937/AAK01938/AAK01939/AAK01940/AAK01941/AAK01942/AAK01943/AAK01944/AAK01945/AAK01946/AAK01947/AAK01948/AAK01949/AAK01950/AAK01951/AAK01952/AAK01953/AAK01954/AAK01955/AAK01956/AAK01957/AAK01958/AAK01959/AAK01960/AAK01961/AAK01962/AAK01963/AAK01964/AAK01965/AAK01966/AAK01967/AAK01968/AAK01969/AAK01970/AAK01971/AAK01972/AAK01973/AAK01974/AAK01975/AAK01976/AAK01977/AAK01978/AAK01979/AAK01980/AAK01981/AAK01982/AAK01983/AAK01984/AAK01985/AAK01986/AAK01987/AAK01988/AAK01989/AAK01990/AAK01991/AAK01992/AAK01993/AAK01994/AAK01995/AAK01996/AAK01997/AAK01998/AAK01999/AAK02000/AAK02001/AAK02002/AAK02003/AAK02004/AAK02005/AAK02006/AAK02007/AAK02008/AAK02009/AAK02010/AAK02011/AAK02012/AAK02013/AAK02014/AAK02015/AAK02016/AAK02017/AAK02018/AAK02019/AAK02020/AAK02021/AAK02022/AAK02023/AAK02024/AAK02025/AAK02026/AAK02027/AAK02028/AAK02029/AAK02030/AAK02031/AAK02032/AAK02033/AAK02034/AAK02035/AAK02036/AAK02037/AAK02038/AAK02039/AAK02040/AAK02041/AAK02042/AAK02043/AAK02044/AAK02045/AAK02046/AAK02047/AAK02048/AAK02049/AAK02050/AAK02051/AAK02052/AAK02053/AAK02054/AAK02055/AAK02056/AAK02057/AAK02058/AAK02059/AAK02060/AAK02061/AAK02062/AAK02063/AAK02064/AAK02065/AAK02066/AAK02067/AAK02068/AAK02069/AAK02070/AAK02071/AAK02072/AAK02073/AAK02074/AAK02075/AAK02076/AAK02077/AAK02078/AAK02079/AAK02080/AAK02081/AAK02082/AAK02083/AAK02084/AAK02085/AAK02086/AAK02087/AAK02088/AAK02089/AAK02090/AAK02091/AAK02092/AAK02093/AAK02094/AAK02095/AAK02096/AAK02097/AAK02098/AAK02099/AAK02100/AAK02101/AAK02102/AAK02103/AAK02104/AAK02105/AAK02106/AAK02107/AAK02108/AAK02109/AAK02110/AAK02111/AAK02112/AAK02113/AAK02114/AAK02115/AAK02116/AAK02117/AAK02118/AAK02119/AAK02120/AAK02121/AAK02122/AAK02123/AAK02124/AAK02125/AAK02126/AAK02127/AAK02128/AAK02129/AAK02130/AAK02131/AAK02132/AAK02133/AAK02134/AAK02135/AAK02136/AAK02137/AAK02138/AAK02139/AAK02140/AAK02141/AAK02142/AAK02143/AAK02144/AAK02145/AAK02146/AAK02147/AAK02148/AAK02149/AAK02150/AAK02151/AAK02152/AAK02153/AAK02154/AAK02155/AAK02156/AAK02157/AAK02158/AAK02159/AAK02160/AAK02161/AAK02162/AAK02163/AAK02164/AAK02165/AAK02166/AAK02167/AAK02168/AAK02169/AAK02170/AAK02171/AAK02172/AAK02173/AAK02174/AAK02175/AAK02176/AAK02177/AAK02178/AAK02179/AAK02180/AAK02181/AAK02182/AAK02183/AAK02184/AAK02185/AAK02186/AAK02187/AAK02188/AAK02189/AAK02190/AAK02191/AAK02192/AAK02193/AAK02194/AAK02195/AAK02196/AAK02197/AAK02198/AAK02199/AAK02200/AAK02201/AAK02202/AAK02203/AAK02204/AAK02205/AAK02206/AAK02207/AAK02208/AAK02209/AAK02210/AAK02211/AAK02212/AAK02213/AAK02214/AAK02215/AAK02216/AAK02217/AAK02218/AAK02219/AAK02220/AAK02221/AAK02222/AAK02223/AAK02224/AAK02225/AAK02226/AAK02227/AAK02228/AAK02229/AAK02230/AAK02231/AAK02232/AAK02233/AAK02234/AAK02235/AAK02236/AAK02237/AAK02238/AAK02239/AAK02240/AAK02241/AAK02242/AAK02243/AAK02244/AAK02245/AAK02246/AAK02247/AAK02248/AAK02249/AAK02250/AAK02251/AAK02252/AAK02253/AAK02254/AAK02255/AAK02256/AAK02257/AAK02258/AAK02259/AAK02260/AAK02261/AAK02262/AAK02263/AAK02264/AAK02265/AAK02266/AAK02267/AAK02268/AAK02269/AAK02270/AAK02271/AAK02272/AAK02273/AAK02274/AAK02275/AAK02276/AAK02277/AAK02278/AAK02279/AAK02280/AAK02281/AAK02282/AAK02283/AAK02284/AAK02285/AAK02286/AAK02287/AAK02288/AAK02289/AAK02290/AAK02291/AAK02292/AAK02293/AAK02294/AAK02295/AAK02296/AAK02297/AAK02298/AAK02299/AAK02300/AAK02301/AAK02302/AAK02303/AAK02304/AAK02305/AAK02306/AAK02307/AAK02308/AAK02309/AAK02310/AAK02311/AAK02312/AAK02313/AAK02314/AAK02315/AAK02316/AAK02317/AAK02318/AAK02319/AAK02320/AAK02321/AAK02322/AAK02323/AAK02324/AAK02325/AAK02326/AAK02327/AAK02328/AAK02329/AAK02330/AAK02331/AAK02332/AAK02333/AAK02334/AAK02335/AAK02336/AAK02337/AAK02338/AAK02339/AAK02340/AAK02341/AAK02342/AAK02343/AAK02344/AAK02345/AAK02346/AAK02347/AAK02348/AAK02349/AAK02350/AAK02351/AAK02352/AAK02353/AAK02354/AAK02355/AAK02356/AAK02357/AAK02358/AAK02359/AAK02360/AAK02361/AAK02362/AAK02363/AAK02364/AAK02365/AAK02366/AAK02367/AAK02368/AAK02369/AAK02370/AAK02371/AAK02372/AAK02373/AAK02374/AAK02375/AAK02376/AAK02377/AAK02378/AAK02379/AAK02380/AAK02381/AAK02382/AAK02383/AAK02384/AAK02385/AAK02386/AAK02387/AAK02388/AAK02389/AAK02390/AAK02391/AAK02392/AAK02393/AAK02394/AAK02395/AAK02396/AAK02397/AAK02398/AAK02399/AAK02400/AAK02401/AAK02402/AAK02403/AAK02404/AAK02405/AAK02406/AAK02407/AAK02408/AAK02409/AAK02410/AAK02411/AAK02412/AAK02413/AAK02414/AAK02415/AAK02416/AAK02417/AAK02418/AAK02419/AAK02420/AAK02421/AAK02422/AAK02423/AAK02424/AAK02425/AAK02426/AAK02427/AAK02428/AAK02429/AAK02430/AAK02431/AAK02432/AAK02433/AAK02434/AAK02435/AAK02436/AAK02437/AAK02438/AAK02439/AAK02440/AAK02441/AAK02442/AAK02443/AAK02444/AAK02445/AAK02446/AAK02447/AAK02448/AAK02449/AAK02450/AAK02451/AAK02452/AAK02453/AAK02454/AAK02455/AAK02456/AAK02457/AAK02458/AAK02459/AAK02460/AAK02461/AAK02462/AAK02463/AAK02464/AAK02465/AAK02466/AAK02467/AAK02468/AAK02469/AAK02470/AAK02471/AAK02472/AAK02473/AAK02474/AAK02475/AAK02476/AAK02477/AAK02478/AAK02479/AAK02480/AAK02481/AAK02482/AAK02483/AAK02484/AAK02485/AAK02486/AAK02487/AAK02488/AAK02489/AAK02490/AAK02491/AAK02492/AAK02493/AAK02494/AAK02495/AAK02496/AAK02497/AAK02498/AAK02499/AAK02500/AAK02501/AAK02502/AAK02503/AAK02504/AAK02505/AAK02506/AAK02507/AAK02508/AAK02509/AAK02510/AAK02511/AAK02512/AAK02513/AAK02514/AAK02515/AAK02516/AAK02517/AAK02518/AAK02519/AAK02520/AAK02521/AAK02522/AAK02523/AAK02524/AAK02525/AAK02526/AAK02527/AAK02528/AAK02529/AAK02530/AAK02531/AAK02532/AAK02533/AAK02534/AAK02535/AAK02536/AAK02537/AAK02538/AAK02539/AAK02540/AAK02541/AAK02542/AAK02543/AAK02544/AAK02545/AAK02546/AAK02547/AAK02548/AAK02549/AAK02550/AAK02551/AAK02552/AAK02553/AAK02554/AAK02555/AAK02556/AAK02557/AAK02558/AAK02559/AAK02560/AAK02561/AAK02562/AAK02563/AAK02564/AAK02565/AAK02566/AAK02567/AAK02568/AAK02569/AAK02570/AAK02571/AAK02572/AAK02573/AAK02574/AAK02575/AAK02576/AAK02577/AAK02578/AAK02579/AAK02580/AAK02581/AAK02582/AAK02583/AAK02584/AAK02585/AAK02586/AAK02587/AAK02588/AAK02589/AAK02590/AAK02591/AAK02592/AAK02593/AAK02594/AAK02595/AAK02596/AAK02597/AAK02598/AAK02599/AAK02600/AAK02601/AAK02602/AAK02603/AAK02604/AAK02605/AAK02606/AAK02607/AAK02608/AA


```
FT STRAND 46 48
FT TURN 50 51
FT STRAND 54 56
FT HELIX 57 59
SQ SEQUENCE 63 AA; 6862 MW; 5DC10DE75B375F16 CRC64;

Query Match 19.0%; Score 92.5; DB 1; Length 63;
Best Local Similarity 32.8%; Pred. No. 0.0076;
Matches 22; Conservative 5; Mismatches 27; Indels 13; Gaps 4;

QY 6 CGENEKDYSCGKEDCKKYDCVEREDDEPNVCLVRVCHQDCVCE--EGFYRNKDK 63
Db 5 CGNEVWTECTG--CEMKC-----GPDE-----NTPCLMCRPSCSPGGRMTNDGK 53

QY 64 CVAEDC 70
Db 54 CIPASQC 60

RESULT 7
LRP CAEL STANDARD; PRT; 4753 AA.
AC Q04833;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Low-density lipoprotein receptor-related protein precursor (LRP).
GN LRP-1 OR F29D11.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93281621; PubMed=8506301;
RA Yochem J., Greenwald I.;
RT "A gene for a low density lipoprotein receptor-related protein in the
RT nematode Caenorhabditis elegans.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:4572-4576(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilkinson J.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY ACT AS A RECEPTOR FOR THE ENDOCYTOSIS OF
CC EXTRACELLULAR LIGANDS SUCH AS CHYLOMICRON REMNANTS, PROTEASE-
CC INHIBITOR COMPLEXES AND VITELLOGENIN.
CC -!- SUBCELLULAR LOCATION: type I membrane protein.
CC -!- SIMILARITY: Contains 35 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 17 EGF-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M96150; AAA28105.1; -
CC DR EMBL; 273907; CAA98124.1; -
CC DR PIR; A47437; A47437.
CC DR HSP; Q07954; 1CR8.
CC DR WormPep; F29D11.1; CE05765.
CC DR InterPro; IPR000152; ASX_hydroxyl.
CC DR InterPro; IPR001881; EGF_Ca.
CC DR InterPro; IPR006209; EGF_Like.
CC DR InterPro; IPR002172; LDL_receptor_A.
CC DR InterPro; IPR000033; LDL_receptor_rep.
CC DR Pfam; PF00057; ldl_recept_a; 34.
CC DR Pfam; PF00058; ldl_recept_b; 26.
CC DR PRINTS; PR00261; LDLRECEPTOR.
CC DR SMART; SM00179; EGF_CA; 2.
```

```
DR SMART; SM00192; LDLa; 35.
DR SMART; SM00135; LY; 32.
DR PROSITE; PS00010; ASX_HYDROXYL; 6.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 3.
DR PROSITE; PS01209; LDLRA_1; 27.
DR PROSITE; PS00068; LDLRA_2; 34.
KW Receptor; Transmembrane; Repeat; Endocytosis; Glycoprotein;
KW Signal; Calcium-binding; EGF-like domain; Coated pits.
FT SIGNAL 1 18
FT CHAIN 19 4753
FT DOMAIN 19 4570
FT TRANSMEM 4571 4596
FT DOMAIN 4597 4753
FT DOMAIN 51 89
FT DOMAIN 90 133
FT DOMAIN 136 177
FT DOMAIN 180 220
FT DOMAIN 221 259
FT DOMAIN 260 298
FT DOMAIN 299 337
FT DOMAIN 338 368
FT DOMAIN 669 712
FT DOMAIN 997 1043
FT DOMAIN 1052 1097
FT DOMAIN 1099 1140
FT DOMAIN 1144 1184
FT DOMAIN 1185 1225
FT DOMAIN 1226 1265
FT DOMAIN 1268 1309
FT DOMAIN 1311 1352
FT DOMAIN 1357 1397
FT DOMAIN 1398 1436
FT DOMAIN 1437 1476
FT DOMAIN 1747 1786
FT DOMAIN 2080 2120
FT DOMAIN 2120 2139
FT DOMAIN 2139 2158
FT DOMAIN 2158 2177
FT DOMAIN 2177 2196
FT DOMAIN 2196 2215
FT DOMAIN 2215 2234
FT DOMAIN 2234 2253
FT DOMAIN 2253 2272
FT DOMAIN 2272 2291
FT DOMAIN 2291 2310
FT DOMAIN 2310 2329
FT DOMAIN 2329 2348
FT DOMAIN 2348 2367
FT DOMAIN 2367 2386
FT DOMAIN 2386 2405
FT DOMAIN 2405 2424
FT DOMAIN 2424 2443
FT DOMAIN 2443 2462
FT DOMAIN 2462 2481
FT DOMAIN 2481 2500
FT DOMAIN 2500 2519
FT DOMAIN 2519 2538
FT DOMAIN 2538 2557
FT DOMAIN 2557 2576
FT DOMAIN 2576 2595
FT DOMAIN 2595 2614
FT DOMAIN 2614 2633
FT DOMAIN 2633 2652
FT DOMAIN 2652 2671
FT DOMAIN 2671 2690
FT DOMAIN 2690 2709
FT DOMAIN 2709 2728
FT DOMAIN 2728 2747
FT DOMAIN 2747 2766
FT DOMAIN 2766 2785
FT DOMAIN 2785 2804
FT DOMAIN 2804 2823
FT DOMAIN 2823 2842
FT DOMAIN 2842 2861
FT DOMAIN 2861 2880
FT DOMAIN 2880 2899
FT DOMAIN 2899 2918
FT DOMAIN 2918 2937
FT DOMAIN 2937 2956
FT DOMAIN 2956 2975
FT DOMAIN 2975 2994
FT DOMAIN 2994 3013
FT DOMAIN 3013 3032
FT DOMAIN 3032 3051
FT DOMAIN 3051 3070
FT DOMAIN 3070 3089
FT DOMAIN 3089 3108
FT DOMAIN 3108 3127
FT DOMAIN 3127 3146
FT DOMAIN 3146 3165
FT DOMAIN 3165 3184
FT DOMAIN 3184 3203
FT DOMAIN 3203 3222
FT DOMAIN 3222 3241
FT DOMAIN 3241 3260
FT DOMAIN 3260 3279
FT DOMAIN 3279 3298
FT DOMAIN 3298 3317
FT DOMAIN 3317 3336
FT DOMAIN 3336 3355
FT DOMAIN 3355 3374
FT DOMAIN 3374 3393
FT DOMAIN 3393 3412
FT DOMAIN 3412 3431
FT DOMAIN 3431 3450
FT DOMAIN 3450 3469
FT DOMAIN 3469 3488
FT DOMAIN 3488 3507
FT DOMAIN 3507 3526
FT DOMAIN 3526 3545
FT DOMAIN 3545 3564
FT DOMAIN 3564 3583
FT DOMAIN 3583 3602
FT DOMAIN 3602 3621
FT DOMAIN 3621 3640
FT DOMAIN 3640 3659
FT DOMAIN 3659 3678
FT DOMAIN 3678 3697
FT DOMAIN 3697 3716
FT DOMAIN 3716 3735
FT DOMAIN 3735 3754
FT DOMAIN 3754 3773
FT DOMAIN 3773 3792
FT DOMAIN 3792 3811
FT DOMAIN 3811 3830
FT DOMAIN 3830 3849
FT DOMAIN 3849 3868
FT DOMAIN 3868 3887
FT DOMAIN 3887 3906
FT DOMAIN 3906 3925
FT DOMAIN 3925 3944
FT DOMAIN 3944 3963
FT DOMAIN 3963 3982
FT DOMAIN 3982 4001
FT DOMAIN 4001 4020
FT DOMAIN 4020 4039
FT DOMAIN 4039 4058
FT DOMAIN 4058 4077
FT DOMAIN 4077 4096
FT DOMAIN 4096 4115
FT DOMAIN 4115 4134
FT DOMAIN 4134 4153
FT DOMAIN 4153 4172
FT DOMAIN 4172 4191
FT DOMAIN 4191 4210
FT DOMAIN 4210 4229
FT DOMAIN 4229 4248
FT DOMAIN 4248 4267
FT DOMAIN 4267 4286
FT DOMAIN 4286 4305
FT DOMAIN 4305 4324
FT DOMAIN 4324 4343
FT DOMAIN 4343 4362
FT DOMAIN 4362 4381
FT DOMAIN 4381 4400
FT DOMAIN 4400 4419
FT DOMAIN 4419 4438
FT DOMAIN 4438 4457
FT DOMAIN 4457 4476
FT DOMAIN 4476 4495
FT DOMAIN 4495 4514
FT DOMAIN 4514 4533
FT DOMAIN 4533 4552
FT DOMAIN 4552 4571
FT DOMAIN 4571 4590
FT DOMAIN 4590 4609
FT DOMAIN 4609 4628
FT DOMAIN 4628 4647
FT DOMAIN 4647 4666
FT DOMAIN 4666 4685
FT DOMAIN 4685 4704
FT DOMAIN 4704 4723
FT DOMAIN 4723 4742
FT DOMAIN 4742 4761
FT DOMAIN 4761 4780
FT DOMAIN 4780 4799
FT DOMAIN 4799 4818
FT DOMAIN 4818 4837
FT DOMAIN 4837 4856
FT DOMAIN 4856 4875
FT DOMAIN 4875 4894
FT DOMAIN 4894 4913
FT DOMAIN 4913 4932
FT DOMAIN 4932 4951
FT DOMAIN 4951 4970
FT DOMAIN 4970 4989
FT DOMAIN 4989 5008
FT DOMAIN 5008 5027
FT DOMAIN 5027 5046
FT DOMAIN 5046 5065
FT DOMAIN 5065 5084
FT DOMAIN 5084 5103
FT DOMAIN 5103 5122
FT DOMAIN 5122 5141
FT DOMAIN 5141 5160
FT DOMAIN 5160 5179
FT DOMAIN 5179 5198
FT DOMAIN 5198 5217
FT DOMAIN 5217 5236
FT DOMAIN 5236 5255
FT DOMAIN 5255 5274
FT DOMAIN 5274 5293
FT DOMAIN 5293 5312
FT DOMAIN 5312 5331
FT DOMAIN 5331 5350
FT DOMAIN 5350 5369
FT DOMAIN 5369 5388
FT DOMAIN 5388 5407
FT DOMAIN 5407 5426
FT DOMAIN 5426 5445
FT DOMAIN 5445 5464
FT DOMAIN 5464 5483
FT DOMAIN 5483 5502
FT DOMAIN 5502 5521
FT DOMAIN 5521 5540
FT DOMAIN 5540 5559
FT DOMAIN 5559 5578
FT DOMAIN 5578 5597
FT DOMAIN 5597 5616
FT DOMAIN 5616 5635
FT DOMAIN 5635 5654
FT DOMAIN 5654 5673
FT DOMAIN 5673 5692
FT DOMAIN 5692 5711
FT DOMAIN 5711 5730
FT DOMAIN 5730 5749
FT DOMAIN 5749 5768
FT DOMAIN 5768 5787
FT DOMAIN 5787 5806
FT DOMAIN 5806 5825
FT DOMAIN 5825 5844
FT DOMAIN 5844 5863
FT DOMAIN 5863 5882
FT DOMAIN 5882 5901
FT DOMAIN 5901 5920
FT DOMAIN 5920 5939
FT DOMAIN 5939 5958
FT DOMAIN 5958 5977
FT DOMAIN 5977 5996
FT DOMAIN 5996 6015
FT DOMAIN 6015 6034
FT DOMAIN 6034 6053
FT DOMAIN 6053 6072
FT DOMAIN 6072 6091
FT DOMAIN 6091 6110
FT DOMAIN 6110 6129
FT DOMAIN 6129 6148
FT DOMAIN 6148 6167
FT DOMAIN 6167 6186
FT DOMAIN 6186 6205
FT DOMAIN 6205 6224
FT DOMAIN 6224 6243
FT DOMAIN 6243 6262
FT DOMAIN 6262 6281
FT DOMAIN 6281 6300
FT DOMAIN 6300 6319
FT DOMAIN 6319 6338
FT DOMAIN 6338 6357
FT DOMAIN 6357 6376
FT DOMAIN 6376 6395
FT DOMAIN 6395 6414
FT DOMAIN 6414 6433
FT DOMAIN 6433 6452
FT DOMAIN 6452 6471
FT DOMAIN 6471 6490
FT DOMAIN 6490 6509
FT DOMAIN 6509 6528
FT DOMAIN 6528 6547
FT DOMAIN 6547 6566
FT DOMAIN 6566 6585
FT DOMAIN 6585 6604
FT DOMAIN 6604 6623
FT DOMAIN 6623 6642
FT DOMAIN 6642 6661
FT DOMAIN 6661 6680
FT DOMAIN 6680 6699
FT DOMAIN 6699 6718
FT DOMAIN 6718 6737
FT DOMAIN 6737 6756
FT DOMAIN 6756 6775
FT DOMAIN 6775 6794
FT DOMAIN 6794 6813
FT DOMAIN 6813 6832
FT DOMAIN 6832 6851
FT DOMAIN 6851 6870
FT DOMAIN 6870 6889
FT DOMAIN 6889 6908
FT DOMAIN 6908 6927
FT DOMAIN 6927 6946
FT DOMAIN 6946 6965
FT DOMAIN 6965 6984
FT DOMAIN 6984 7003
FT DOMAIN 7003 7022
FT DOMAIN 7022 7041
FT DOMAIN 7041 7060
FT DOMAIN 7060 7079
FT DOMAIN 7079 7098
FT DOMAIN 7098 7117
FT DOMAIN 7117 7136
FT DOMAIN 7136 7155
FT DOMAIN 7155 7174
FT DOMAIN 7174 7193
FT DOMAIN 7193 7212
FT DOMAIN 7212 7231
FT DOMAIN 7231 7250
FT DOMAIN 7250 7269
FT DOMAIN 7269 7288
FT DOMAIN 7288 7307
FT DOMAIN 7307 7326
FT DOMAIN 7326 7345
FT DOMAIN 7345 7364
FT DOMAIN 7364 7383
FT DOMAIN 7383 7402
FT DOMAIN 7402 7421
FT DOMAIN 7421 7440
FT DOMAIN 7440 7459
FT DOMAIN 7459 7478
FT DOMAIN 7478 7497
FT DOMAIN 7497 7516
FT DOMAIN 7516 7535
FT DOMAIN 7535 7554
FT DOMAIN 7554 7573
FT DOMAIN 7573 7592
FT DOMAIN 7592 7611
FT DOMAIN 7611 7630
FT DOMAIN 7630 7649
FT DOMAIN 7649 7668
FT DOMAIN 7668 7687
FT DOMAIN 7687 7706
FT DOMAIN 7706 7725
FT DOMAIN 7725 7744
FT DOMAIN 7744 7763
FT DOMAIN 7763 7782
FT DOMAIN 7782 7801
FT DOMAIN 7801 7820
FT DOMAIN 7820 7839
FT DOMAIN 7839 7858
FT DOMAIN 7858 7877
FT DOMAIN 7877 7896
FT DOMAIN 7896 7915
FT DOMAIN 7915 7934
FT DOMAIN 7934 7953
FT DOMAIN 7953 7972
FT DOMAIN 7972 7991
FT DOMAIN 7991 8010
FT DOMAIN 8010 8029
FT DOMAIN 8029 8048
FT DOMAIN 8048 8067
FT DOMAIN 8067 8086
FT DOMAIN 8086 8105
FT DOMAIN 8105 8124
FT DOMAIN 8124 8143
FT DOMAIN 8143 8162
FT DOMAIN 8162 8181
FT DOMAIN 8181 8200
FT DOMAIN 8200 8219
FT DOMAIN 8219 8238
FT DOMAIN 8238 8257
FT DOMAIN 8257 8276
FT DOMAIN 8276 8295
FT DOMAIN 8295 8314
FT DOMAIN 8314 8333
FT DOMAIN 8333 8352
FT DOMAIN 8352 8371
FT DOMAIN 8371 8390
FT DOMAIN 8390 8409
FT DOMAIN 8409 8428
FT DOMAIN 8428 8447
FT DOMAIN 8447 8466
FT DOMAIN 8466 8485
FT DOMAIN 8485 8504
FT DOMAIN 8504 8523
FT DOMAIN 8523 8542
FT DOMAIN 8542 8561
FT DOMAIN 8561 8580
FT DOMAIN 8580 8599
FT DOMAIN 8599 8618
FT DOMAIN 8618 8637
FT DOMAIN 8637 8656
FT DOMAIN 8656 8675
FT DOMAIN 8675 8694
FT DOMAIN 8694 8713
FT DOMAIN 8713 8732
FT DOMAIN 8732 8751
FT DOMAIN 8751 8770
FT DOMAIN 8770 8789
FT DOMAIN 8789 8808
FT DOMAIN 8808 8827
FT DOMAIN 8827 8846
FT DOMAIN 8846 8865
FT DOMAIN 8865 8884
FT DOMAIN 8884 8903
FT DOMAIN 8903 8922
FT DOMAIN 8922 8941
FT DOMAIN 8941 8960
FT DOMAIN 8960 8979
FT DOMAIN 8979 8998
FT DOMAIN 8998 9017
FT DOMAIN 9017 9036
FT DOMAIN 9036 9055
FT DOMAIN 9055 9074
FT DOMAIN 9074 9093
FT DOMAIN 9093 9112
FT DOMAIN 9112 9131
FT DOMAIN 9131 9150
FT DOMAIN 9150 9169
FT DOMAIN 9169 9188
FT DOMAIN 9188 9207
FT DOMAIN 9207 9226
FT DOMAIN 9226 9245
FT DOMAIN 9245 9264
FT DOMAIN 9264 9283
FT DOMAIN 9283 9302
FT DOMAIN 9302 9321
FT DOMAIN 9321 9340
FT DOMAIN 9340 9359
FT DOMAIN 9359 9378
FT DOMAIN 9378 9397
FT DOMAIN 9397 9416
FT DOMAIN 9416 9435
FT DOMAIN 9435 9454
FT DOMAIN 9454 9473
FT DOMAIN 9473 9492
FT DOMAIN 9492 9511
FT DOMAIN 9511 9530
FT DOMAIN 9530 9549
FT DOMAIN 9549 9568
FT DOMAIN 9568 9587
FT DOMAIN 9587 9606
FT DOMAIN 9606 9625
FT DOMAIN 9625 9644
FT DOMAIN 9644 9663
FT DOMAIN 9663 9682
FT DOMAIN 9682 9701
FT DOMAIN 9701 9720
FT DOMAIN 9720 9739
FT DOMAIN 9739 9758
FT DOMAIN 9758 9777
FT DOMAIN 9777 9796
FT DOMAIN 9796 9815
FT DOMAIN 9815 9834
FT DOMAIN 9834 9853
FT DOMAIN 9853 9872
FT DOMAIN 9872 9891
FT DOMAIN 9891 9910
FT DOMAIN 9910 9929
FT DOMAIN 9929 9948
FT DOMAIN 9948 9967
FT DOMAIN 9967 9986
FT DOMAIN 9986 10005
FT DOMAIN 10005 10024
FT DOMAIN 10024 10043
FT DOMAIN 10043 10062
FT DOMAIN 10062 10081
FT DOMAIN 10081 10100
FT DOMAIN 10100 10119
FT DOMAIN 10119 10138
FT DOMAIN 10138 10157
FT DOMAIN 10157 10176
FT DOMAIN 10176 10195
FT DOMAIN 10195 10214
FT DOMAIN 10214 10233
FT DOMAIN 10233 10252
FT DOMAIN 10252 10271
FT DOMAIN 10271 10290
FT DOMAIN 10290 10309
FT DOMAIN 10309 10328
FT DOMAIN 10328 10347
FT DOMAIN 10347 10366
FT DOMAIN 10366 10385
FT DOMAIN 10385 10404
FT DOMAIN 10404 10423
FT DOMAIN 10423 10442
FT DOMAIN 10442 10461
FT DOMAIN 10461 10480
FT DOMAIN 10480 10499
FT DOMAIN 10499 10518
FT DOMAIN 10518 10537
FT DOMAIN 10537 10556
FT DOMAIN 10556 10575
FT DOMAIN 10575 10594
FT DOMAIN 10594 10613
FT DOMAIN 10613 10632
FT DOMAIN 10632 10651
FT DOMAIN 10651 10670
FT DOMAIN 10670 10689
FT DOMAIN 10689 10708
FT DOMAIN 10708 10727
FT DOMAIN 10727 10746
FT DOMAIN 10746 10765
FT DOMAIN 10765 10784
FT DOMAIN 10784 10803
FT DOMAIN 10803 10822
FT DOMAIN 10822 10841
FT DOMAIN 10841 10860
FT DOMAIN 10860 10879
FT DOMAIN 10879 10898
FT DOMAIN 10898 10917
FT DOMAIN 10917 10936
FT DOMAIN 10936 10955
FT DOMAIN 10955 10974
FT DOMAIN 10974 10993
FT DOMAIN 10993 11012
FT DOMAIN 11012 11031
FT DOMAIN 11031 11050
FT DOMAIN 11050 11069
FT DOMAIN 11069 11088
FT DOMAIN 11088 11107
FT DOMAIN 11107 11126
FT DOMAIN 11126 11145
FT DOMAIN 11145 11164
FT DOMAIN 11164 11183
FT DOMAIN 11183 11202
FT DOMAIN 11202 11221
FT DOMAIN 11221 11240
FT DOMAIN 11240 11259
FT DOMAIN 11259 11278
FT DOMAIN 11278 11297
FT DOMAIN 11297 11316
FT DOMAIN 11316 11335
FT DOMAIN 11335 11354
FT DOMAIN 11354 11373
FT DOMAIN 11373 11392
FT DOMAIN 11392 11411
FT DOMAIN 11411 11430
FT DOMAIN 11430 11449
FT DOMAIN 11449 11468
FT DOMAIN 11468 11487
FT DOMAIN 11487 11506
FT DOMAIN 11506 11525
FT DOMAIN 11525 11544
FT DOMAIN 11544 11563
FT DOMAIN 11563 11582
FT DOMAIN 11582 11601
FT DOMAIN 11601 11620
FT DOMAIN 11620 11639
FT DOMAIN 11639 11658
FT DOMAIN 11658 11677
FT DOMAIN 11677 11696
FT DOMAIN 11696 11715
FT DOMAIN 11715 11734
FT DOMAIN 11734 11753
FT DOMAIN 11753 11772
FT DOMAIN 11772 11791
FT DOMAIN 11791 11810
FT DOMAIN 11810 11829
FT DOMAIN 11829 11848
FT DOMAIN 11848 11867
FT DOMAIN 11867 11886
FT DOMAIN 11886 11905
FT DOMAIN 11905 11924
FT DOMAIN 11924 11943
FT DOMAIN 11943 11962
FT DOMAIN 11962 11981
FT DOMAIN 11981 12000
FT DOMAIN 12000 12019
FT DOMAIN 12019 12038
FT DOMAIN 12038 12057
FT DOMAIN 12057 12076
FT DOMAIN 12076 12095
FT DOMAIN 12095 12114
FT DOMAIN 12114 12133
FT DOMAIN 12133 12152
FT DOMAIN 12152 12171
FT DOMAIN 12171 12190
FT DOMAIN 12190 12209
FT DOMAIN 12209 12228
FT DOMAIN 12228 12247
FT DOMAIN 12247 12266
FT DOMAIN 12266 12285
FT DOMAIN 12285 12304
FT DOMAIN 12304 12323
FT DOMAIN 12323 12342
FT DOMAIN 12342 12361
FT DOMAIN 12361 12380
FT DOMAIN 12380 12399
FT DOMAIN 12399 12418
FT DOMAIN 12418 12437
FT DOMAIN 12437 12456
FT DOMAIN 12456 12475
FT DOMAIN 12475 12494
FT DOMAIN 12494 12513
FT DOMAIN 12513 12532
FT DOMAIN 12532 12551
FT DOMAIN 12551 12570
FT DOMAIN 12570 12589
FT DOMAIN 12589 12608
FT DOMAIN 12608 12627
FT DOMAIN 12627 12646
FT DOMAIN 12646 12665
FT DOMAIN 12665 12684
FT DOMAIN 12684 12703
FT DOMAIN 12703 12722
FT DOMAIN 12722 12741
FT DOMAIN 12741 12760
FT DOMAIN 12760 12779
FT DOMAIN 12779 12798
FT DOMAIN 12798 12817
FT DOMAIN 12817 12836
FT DOMAIN 12836 12855
FT DOMAIN 12855 12874
FT DOMAIN 12874 12893
FT DOMAIN 12893 12912
FT DOMAIN 12912 12931
FT DOMAIN 12931 12950
FT DOMAIN 12950 12969
FT DOMAIN 12969 12988
FT DOMAIN 12988 13007
FT DOMAIN 13007 13026
FT DOMAIN 13026 13045
FT DOMAIN 13045 13064
FT DOMAIN 13064 13083
FT DOMAIN 13083 13102
FT DOMAIN 13102 13121
FT DOMAIN 13121 13140
FT DOMAIN 13140 13159
FT DOMAIN 13159 13178
FT DOMAIN 13178 13197
FT DOMAIN 13197 13216
FT DOMAIN 13216 13235
FT DOMAIN 13235 13254
FT DOMAIN 13254 13273
FT DOMAIN 13273 13292
FT DOMAIN 13292 13311
FT DOMAIN 13311 13330
FT DOMAIN 13330 13349
FT DOMAIN 13349 13368
FT DOMAIN 13368 13387
FT DOMAIN 13387 13406
FT DOMAIN 13406 13425
FT DOMAIN 13425 13444
FT DOMAIN 13444 13463
FT DOMAIN 13463 13482
FT DOMAIN 13482 13501
FT DOMAIN 13501 13520
FT DOMAIN 13520 13539
FT DOMAIN 13539 13558
FT DOMAIN 13558 13577
FT DOMAIN 13577 13596
FT DOMAIN 13596 13615
FT DOMAIN 13615 13634
FT DOMAIN 13634 13653
FT DOMAIN 13653 13672
FT DOMAIN 13672 13691
FT DOMAIN 13691 13710
FT DOMAIN 13710 13729
FT DOMAIN 13729 13748
FT DOMAIN 13748 13767
FT DOMAIN 13767 13786
FT DOMAIN 13786 13805
FT DOMAIN 13805 13824
FT DOMAIN 13824 13843
FT DOMAIN 13843 13862
FT DOMAIN 13862 13881
FT DOMAIN 13881 13900
FT DOMAIN 13900 13919
FT DOMAIN 13919 13938
FT DOMAIN 13938 13957
FT DOMAIN 13957 13976
FT DOMAIN 13976 13995
FT DOMAIN 13995 14014
FT DOMAIN 14014 14033
FT DOMAIN 14033 14052
FT DOMAIN 14052 14071
FT DOMAIN 14071 14090
FT DOMAIN 14090 14109
FT DOMAIN 14109 14128
FT DOMAIN 14128 14147
FT DOMAIN 14147 14166
FT DOMAIN 14166 14185
FT DOMAIN 14185 14204
FT DOMAIN 14204 14223
FT DOMAIN 14223 14242
FT DOMAIN 14242 14261
FT DOMAIN 14261 14280
FT DOMAIN 14280 14299
FT DOMAIN 14299 14318
FT DOMAIN 14318 14337
FT DOMAIN 14337 14356
FT DOMAIN 14356 14375
FT DOMAIN 14375 14394
FT DOMAIN 14394 14413
FT DOMAIN 14413 14432
FT DOMAIN 14432 14451
FT DOMAIN 14451 14470
FT DOMAIN 14470 14489
FT DOMAIN 14489 14508
FT DOMAIN 14508 14527
FT DOMAIN 14527 14546
FT DOMAIN 14546 14565
FT DOMAIN 14565 14584
FT DOMAIN 14584 14603
FT DOMAIN 14603 14622
FT DOMAIN 14622 14641
FT DOMAIN 14641 14660
FT DOMAIN 14660 14679
FT DOMAIN 14679 14698
FT DOMAIN 14698 14717
FT DOMAIN 14717 14736
FT DOMAIN 14736 14755
FT DOMAIN 14755 14774
FT DOMAIN 14774 14793
FT DOMAIN 14793 14812
FT DOMAIN 14812 14831
FT DOMAIN 14831 14850
FT DOMAIN 14850 14869
FT DOMAIN 14869 14888
FT DOMAIN 14888 14907
FT DOMAIN 14907 14926
FT DOMAIN 14926 14945
FT DOMAIN 14945 14964
FT DOMAIN 14964 14983
FT DOMAIN 14983 15002
FT DOMAIN 15002 15021
FT DOMAIN 15021 15040
FT DOMAIN 15040 15059
FT DOMAIN 15059 15078
FT DOMAIN 15078 15097
FT DOMAIN 15097 15116
FT DOMAIN 15116 15135
FT DOMAIN 15135 15154
FT DOMAIN 15154 15173
FT DOMAIN 15173 15192
FT DOMAIN 15192 15211
FT DOMAIN 15211 15230
FT DOMAIN 15230 15249
FT DOMAIN 15249 15268
FT DOMAIN 15268 15287
FT DOMAIN 15287 15306
FT DOMAIN 15306 15325
FT DOMAIN 15325 15344
FT DOMAIN 15344 15363
FT DOMAIN 15363 15382
FT DOMAIN 15382 15401
FT DOMAIN 15401 15420
FT DOMAIN 15420 15439
FT DOMAIN 15439 15458
FT DOMAIN 15458 15477
FT DOMAIN 15477 15496
FT DOMAIN 15496 15515
FT DOMAIN 15515 15534
FT DOMAIN 15534 15553
FT DOMAIN 15553 15572
FT DOMAIN 15572 15591
FT DOMAIN 15591 15610
FT DOMAIN 15610 15629
FT DOMAIN 15629 15648
FT DOMAIN 15648 15667
FT DOMAIN 15667 15686
FT DOMAIN 15686 15705
FT DOMAIN 15705 15724
FT DOMAIN 15724 15743
FT DOMAIN 15743 15762
FT DOMAIN 15762 15781
FT DOMAIN 15781 15800
FT DOMAIN 15800 15819
FT DOMAIN 15819 15838
FT DOMAIN 15838 15857
FT DOMAIN 15857 15876
FT DOMAIN 15876 15895
FT DOMAIN 15895 15914
FT DOMAIN 15914 15933
FT DOMAIN 15933 15952
FT DOMAIN 15952 15971
FT DOMAIN 15971 15990
FT DOMAIN 15990 16009
FT DOMAIN 16009 16028
FT DOMAIN 16028 16047
FT DOMAIN 16047 16066
FT DOMAIN 16066 16085
FT DOMAIN 16085 16104
FT DOMAIN 16104 16123
FT DOMAIN 16123 16142
FT DOMAIN 16142 16161
FT DOMAIN 16161 16180
FT DOMAIN 16180 16199
FT DOMAIN 16199 16218
FT DOMAIN 16218 16237
FT DOMAIN 16237 16256
FT DOMAIN 16256 16275
FT DOMAIN 16275 16294
FT DOMAIN 16294 16313
FT DOMAIN 16313 16332
FT DOMAIN 16332 16351
FT DOMAIN 16351 16370
FT DOMAIN 16370 16389
FT DOMAIN 16389 16408
FT DOMAIN 16408 16427
FT DOMAIN 16427 16446
FT DOMAIN 16446 16465
FT DOMAIN 16465
```

```
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 99 122 BY SIMILARITY.
FT DISULFID 116 131 BY SIMILARITY.
FT DISULFID 138 152 BY SIMILARITY.
FT DISULFID 146 165 BY SIMILARITY.
FT DISULFID 159 175 BY SIMILARITY.
FT DISULFID 182 195 BY SIMILARITY.
FT DISULFID 189 208 BY SIMILARITY.
FT DISULFID 202 218 BY SIMILARITY.
FT DISULFID 223 235 BY SIMILARITY.
FT DISULFID 230 248 BY SIMILARITY.
FT DISULFID 242 257 BY SIMILARITY.
FT DISULFID 262 275 BY SIMILARITY.
FT DISULFID 269 288 BY SIMILARITY.
FT DISULFID 282 297 BY SIMILARITY.
FT DISULFID 302 311 BY SIMILARITY.
FT DISULFID 307 320 BY SIMILARITY.
FT DISULFID 322 336 BY SIMILARITY.
FT DISULFID 342 352 BY SIMILARITY.
FT DISULFID 348 361 BY SIMILARITY.
FT DISULFID 363 367 BY SIMILARITY.
FT DISULFID 673 682 BY SIMILARITY.
FT DISULFID 678 697 BY SIMILARITY.
FT DISULFID 699 711 BY SIMILARITY.
FT DISULFID 1001 1010 BY SIMILARITY.
FT DISULFID 1006 1026 BY SIMILARITY.
FT DISULFID 1028 1042 BY SIMILARITY.
FT DISULFID 1054 1068 BY SIMILARITY.
FT DISULFID 1063 1081 BY SIMILARITY.
FT DISULFID 1075 1095 BY SIMILARITY.
FT DISULFID 1101 1114 BY SIMILARITY.
FT DISULFID 1108 1127 BY SIMILARITY.
FT DISULFID 1121 1138 BY SIMILARITY.
FT DISULFID 1146 1158 BY SIMILARITY.
FT DISULFID 1153 1171 BY SIMILARITY.
FT DISULFID 1165 1182 BY SIMILARITY.
FT DISULFID 1187 1199 BY SIMILARITY.
FT DISULFID 1194 1212 BY SIMILARITY.
FT DISULFID 1206 1223 BY SIMILARITY.
FT DISULFID 1228 1241 BY SIMILARITY.
FT DISULFID 1235 1254 BY SIMILARITY.
FT DISULFID 1248 1263 BY SIMILARITY.
FT DISULFID 1270 1283 BY SIMILARITY.
FT DISULFID 1277 1296 BY SIMILARITY.
FT DISULFID 1290 1307 BY SIMILARITY.
FT DISULFID 1313 1325 BY SIMILARITY.
FT DISULFID 1320 1338 BY SIMILARITY.
FT DISULFID 1332 1350 BY SIMILARITY.
FT DISULFID 1359 1373 BY SIMILARITY.
FT DISULFID 1366 1386 BY SIMILARITY.
FT DISULFID 1380 1396 BY SIMILARITY.
FT DISULFID 1401 1412 BY SIMILARITY.
FT DISULFID 1408 1421 BY SIMILARITY.
FT DISULFID 1423 1435 BY SIMILARITY.
FT DISULFID 1441 1451 BY SIMILARITY.
FT DISULFID 1447 1460 BY SIMILARITY.
FT DISULFID 1462 1475 BY SIMILARITY.
FT DISULFID 1751 1760 BY SIMILARITY.
FT DISULFID 1756 1770 BY SIMILARITY.
FT DISULFID 1772 1785 BY SIMILARITY.
FT DISULFID 2084 2095 BY SIMILARITY.
FT DISULFID 2091 2105 BY SIMILARITY.
FT DISULFID 2107 2119 BY SIMILARITY.
FT DISULFID 2400 2415 BY SIMILARITY.
FT DISULFID 2411 2426 BY SIMILARITY.
FT DISULFID 2428 2438 BY SIMILARITY.
FT DISULFID 2732 2743 BY SIMILARITY.
FT DISULFID 2739 2759 BY SIMILARITY.
FT DISULFID 2761 2779 BY SIMILARITY.
FT DISULFID 2792 2805 BY SIMILARITY.
FT DISULFID 2800 2818 BY SIMILARITY.
FT DISULFID 2812 2829 BY SIMILARITY.
FT DISULFID 2834 2846 BY SIMILARITY.

FT DISULFID 2841 2859 BY SIMILARITY.
FT DISULFID 2853 2868 BY SIMILARITY.

Query Match 19.0%; Score 92.5; DB 1; Length 4753;
Best Local Similarity 25.0%; Pred. No. 0.57;
Matches 28; Conservative 11; Mismatches 32; Indels 41; Gaps 5;

QY 7 GENEKVDSCGSKEC-----DKKCKYGVVEEDDEEVPVPC---- 41
DB 1174 GDEKLEMGNATCAANQFSCANGRCIPYIWLCDGNDNC-YDGTDEKRCPPVQCSALQ 1232
QY 42 -----LVRVCHQDCVCEEGFYRNKDKC-VSAEDCELDNNDFIYPG 81
DB 1233 FRCANGRCQCVPLRNHCDGQSDCEDG---SDEDSCAVTAESCTPDQFKCVSSG 1281

RESULT 8
VWF CANFA
ID VWF CANFA STANDARD; PRT; 2813 AA.
AC Q28295; Q28311; Q9TS14;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Von Willebrand factor precursor (vWF).
GN VWF OR F8VWF.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Stoy S.J., Shibuya H., Nonneman D.J., Holzhauser J., Mohammed I.H.,
RA Johnson G.S.;
RA Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Montgomery R.R., Fahs S., Montgomery M.W.;
RA Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Venta P.J., Li J., Yuzbasiyan-Gurkan V., Brewer G.J., Schall W.D.;
RA "Complete sequence of the structural gene for canine von Willebrand
RT factor and identification of a mutation causing Scottish terrier von
RT Willebrand's disease.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1234-1669 FROM N.A.
RC TISSUE=Blood;
RA Mancuso D.J., Christopherson P.A., Kroner P.A., Montgomery R.R.;
RT "The canine von Willebrand factor gene: sequence and expression of
RT a region encoding the glycoprotein Ib/IX binding domain.";
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: IMPORTANT IN THE MAINTENANCE OF HOMEOSTASIS, IT
CC PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A
CC NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF
CC VASCULAR INJURY (BY SIMILARITY).
CC -I- SUBUNIT: MULTIMERIC (BY SIMILARITY).
CC -I- TISSUE SPECIFICITY: BLOOD.
CC -I- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC -I- SIMILARITY: Contains 3 vWFA domains.
CC -I- SIMILARITY: Contains 3 vWFC domains.
CC -I- SIMILARITY: Contains 4 vWFD domains.
CC -I- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -I- SIMILARITY: SOME, TO SILKWORM HEMOCYTIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
```

```

CC EMBL; L76227; AAB05549.1; -
DR DR EMBL; L16903; AAA30903.1; -
DR DR EMBL; AF099154; AAD04919.1; -
DR DR EMBL; U66246; AAB93766.1; -
DR DR HSSP; P04275; 1AUQ.
DR InterPro; IPR006208; Cys_knot_C.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000004; SapB.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR002035; VWF_A.
DR InterPro; IPR001007; VWF_C.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF01826; TIL; 4.
DR Pfam; PF00092; vwa; 3.
DR Pfam; PF00093; vwc; 3.
DR Pfam; PF00094; vwg; 4.
DR PRINTS; SM00453; VWFADOMAIN.
DR SMART; SM00041; CT; 1.
DR SMART; SM00327; VWA; 3.
DR SMART; SM00214; VWC; 5.
DR SMART; SM00216; VWD; 4.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS0234; VWFA; 3.
DR PROSITE; PS01208; VWFC_1; 3.
DR PROSITE; PS0184; VWFC_2; 3.
KW Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
KW Plasma; Hemostasis; Repeat; Cell adhesion; Signal.
FT SIGNAL 1 22
FT PROPEP 23 763
FT CHAIN 764 2813
FT DOMAIN 35 179
FT DOMAIN 388 541
FT DOMAIN 764 787
FT DOMAIN 788 833
FT DOMAIN 826 853
FT DOMAIN 867 1013
FT DOMAIN 1277 1453
FT DOMAIN 1498 1665
FT DOMAIN 1691 1871
FT DOMAIN 1950 2102
FT DOMAIN 2216 2261
FT DOMAIN 2255 2328
FT DOMAIN 2429 2495
FT DOMAIN 2580 2645
FT DOMAIN 2724 2812
FT SITE 531 533
FT SITE 698 700
FT SITE 2507 2509
FT DISULFID 767 808
FT DISULFID 776 804
FT DISULFID 810 821
FT DISULFID 867 996
FT DISULFID 889 1031
FT DISULFID 898 993
FT DISULFID 914 921
FT DISULFID 1060 1084
FT DISULFID 1071 1111
FT DISULFID 1089 1091
FT DISULFID 1153 1165
FT DISULFID 1149 1169
FT DISULFID 1126 1130
FT DISULFID 1136 1199
FT DISULFID 1234 1237
FT DISULFID 1272 1458
FT DISULFID 1669 1670
FT DISULFID 1686 1872
FT DISULFID 1879 1904
FT DISULFID 1899 1940
FT DISULFID 1972 2123
FT DISULFID 1950 2085
-----
EMBL; L76227; AAB05549.1; -
DR DR EMBL; L16903; AAA30903.1; -
DR DR EMBL; AF099154; AAD04919.1; -
DR DR EMBL; U66246; AAB93766.1; -
DR DR HSSP; P04275; 1AUQ.
DR InterPro; IPR006208; Cys_knot_C.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000004; SapB.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR002035; VWF_A.
DR InterPro; IPR001007; VWF_C.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF01826; TIL; 4.
DR Pfam; PF00092; vwa; 3.
DR Pfam; PF00093; vwc; 3.
DR Pfam; PF00094; vwg; 4.
DR PRINTS; SM00453; VWFADOMAIN.
DR SMART; SM00041; CT; 1.
DR SMART; SM00327; VWA; 3.
DR SMART; SM00214; VWC; 5.
DR SMART; SM00216; VWD; 4.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS0234; VWFA; 3.
DR PROSITE; PS01208; VWFC_1; 3.
DR PROSITE; PS0184; VWFC_2; 3.
KW Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
KW Plasma; Hemostasis; Repeat; Cell adhesion; Signal.
FT SIGNAL 1 22
FT PROPEP 23 763
FT CHAIN 764 2813
FT DOMAIN 35 179
FT DOMAIN 388 541
FT DOMAIN 764 787
FT DOMAIN 788 833
FT DOMAIN 826 853
FT DOMAIN 867 1013
FT DOMAIN 1277 1453
FT DOMAIN 1498 1665
FT DOMAIN 1691 1871
FT DOMAIN 1950 2102
FT DOMAIN 2216 2261
FT DOMAIN 2255 2328
FT DOMAIN 2429 2495
FT DOMAIN 2580 2645
FT DOMAIN 2724 2812
FT SITE 531 533
FT SITE 698 700
FT SITE 2507 2509
FT DISULFID 767 808
FT DISULFID 776 804
FT DISULFID 810 821
FT DISULFID 867 996
FT DISULFID 889 1031
FT DISULFID 898 993
FT DISULFID 914 921
FT DISULFID 1060 1084
FT DISULFID 1071 1111
FT DISULFID 1089 1091
FT DISULFID 1153 1165
FT DISULFID 1149 1169
FT DISULFID 1126 1130
FT DISULFID 1136 1199
FT DISULFID 1234 1237
FT DISULFID 1272 1458
FT DISULFID 1669 1670
FT DISULFID 1686 1872
FT DISULFID 1879 1904
FT DISULFID 1899 1940
FT DISULFID 1972 2123
FT DISULFID 1950 2085
-----
FT DISULFID 1927 2088 BY SIMILARITY.
FT DISULFID 1993 2001 BY SIMILARITY.
FT DISULFID 2724 2774 BY SIMILARITY.
FT DISULFID 2739 2788 BY SIMILARITY.
FT DISULFID 2750 2804 BY SIMILARITY.
FT DISULFID 2754 2806 BY SIMILARITY.
FT DISULFID ? 2811 BY SIMILARITY.
FT CARBOHYD 99 99 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 666 666 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 857 857 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1231 1231 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1515 1515 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1574 1574 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2223 2223 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2290 2290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2357 2357 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2400 2400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2546 2546 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2585 2585 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2790 2790 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 55 55 G -> E (IN REF. 2).
FT CONFLICT 70 70 V -> I (IN REF. 3).
FT CONFLICT 266 266 A -> G (IN REF. 2).
FT CONFLICT 280 280 I -> V (IN REF. 2).
FT CONFLICT 409 411 VCH -> ICQ (IN REF. 2).
FT CONFLICT 994 994 G -> A (IN REF. 1).
FT CONFLICT 1021 1021 F -> L (IN REF. 2).
FT CONFLICT 2381 2381 L -> P (IN REF. 2).
FT CONFLICT 2406 2406 P -> L (IN REF. 2).
SQ SEQUENCE 2813 AA; 309716 MW; 5DP93E1E5E72F80C CRC64;
Query Match 18.0%; Score 87.5; DB 1; Length 2813;
Best Local Similarity 28.4%; Pred. No. 1;
Matches 19; Conservative 9; Mismatches 30; Indels 9; Gaps 2;
QY 4 MOCGENEKYDSCGKECDKCKYDGVREDEDEEPNVCPLVRVCHODCVCEEFGYRNKDDK 63
DB 650 LSCPGQVYLQGT-PCNWTCSRSLSPEDCNE-----VLEGCFPPGLYDGRGD 700
QY 64 CVSAEDC 70
DB 701 CVPKAQ 707
RESULT 9
AG2F DROME STANDARD; PRT; 115 AA.
AC 046202;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Accessory gland protein Acp62F precursor.
GN ACP62F OR CG1262.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.; FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=Canton-S; TISSUE=Male accessory gland;
RX MEDLINE=98135120; PubMed=9474779;
RA Wolfner M.F., Harada H.A., Bertram M.J., Stelick T.J., Kraus K.W.,
RA Kalb J.M., Lung Y.O., Neubaum D.M., Park M., Tram U.K.;
RT "New genes for male accessory gland proteins in Drosophila
RT melanogaster."
RL Insect Biochem. Mol. Biol. 27:825-834(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;

```

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoekins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
[3]
RP SEQUENCE OF 7-111 FROM N.A.
RC STRAIN=ZIM62H-12C, ZIM62H-16C, ZIM62H-28C, ZIM62H-30C, ZIM62H-34C,
RC ZIM62I-5C, ZIM62I-10C, ZIM62I-17C, ZIM62I-18C, and ZIM62I-53C;
RX MEDLINE=20556153; PubMed=11102381;
RA Begun D.J., Whitley P., Todd B.L., Waldrip-Dail H.M., Clark A.G.;
RT "Molecular population genetics of male accessory gland proteins in
RT Drosophila";
RL Genetics 156:1879-1888 (2000).
CC -!- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN
CC MATED FEMALE FLIES. MAY CONTRIBUTE TO THE TOXICITY OF SEMINAL
CC FLUID AND THE DECREASED LIFE-SPAN OF MATED FEMALES. MAY ALSO
CC AFFECT NEUROMUSCULAR EVENTS AFTER MATING CONCERNING SPERM STORAGE
CC AND EGG RELEASE.
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC -!- TISSUE SPECIFICITY: SEMINAL FLUID.
CC -!- SIMILARITY: SOME, TO P.NIGRIVENTER TX2-6.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi-sib.ch/announce/>
CC or send an email to license@ebi-sib.ch).
CC
CC -----
DR EMBL; U85763; AAB96387.1; -;
DR EMBL; AE003475; AAF47683.1; -;
DR EMBL; AY010608; AAG35367.1; -;
DR EMBL; AY010609; AAG35368.1; -;
DR EMBL; AY010610; AAG35369.1; -;
DR EMBL; AY010611; AAG35370.1; -;
DR EMBL; AY010612; AAG35371.1; -;
DR EMBL; AY010613; AAG35372.1; -;
DR EMBL; AY010614; AAG35373.1; -;

DR EMBL; AY010615; AAG35374.1; -;
DR EMBL; AY010616; AAG35375.1; -;
DR EMBL; AY010617; AAG35376.1; -;
DR FlyBase; FBgn0020509; ACP62F.
DR GO; GO:0004867; F:serine protease inhibitor activity; IDA.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
KW Behavior; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 115 ACCESSORY GLAND PROTEIN ACP62F.
FT DOMAIN 34 88 TIL.
SQ SEQUENCE 115 AA; 12570 MW; 4326AA6F6C32291D CRC64;

Query Match 17.9%; Score 87; DB 1; Length 115;
Best Local Similarity 31.0%; Pred. No. 0.047;
Matches 22; Conservative 8; Mismatches 29; Indels 12; Gaps 4;

QY 4 MCGENKDYDSGSKCKKCKYGVVEEDDEPNVPCLVRVCHQDCVCEGFPYRN-KDD 62
Db 32 VDCNTANGTQTEC-PVACPETCEYSG-----NGPC-VKMGAFCKPGVYNERIP 80
QY 63 KCVSAEDCDEL 73
Db 81 ACVLRSDCPKD 91

RESULT 10
LRP2_HUMAN STANDARD; PRT; 4655 AA.
ID LRP2_HUMAN STANDARD; PRT; 4655 AA.
AC P98164; O00711; Q16215;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Low-density lipoprotein receptor-related protein 2 precursor (Megalin)
DE (Glycoprotein 330) (gp330).
GN LRP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=96305376; PubMed=8706697;
RA Hjaelm G., Murray E., Crumley G., Harazim W., Lundgren S., Onyango I.,
RA Ek B., Larsson M., Juhlin C., Hellman P., Davis H., Akerstrom G.,
RA Rask L., Morse B.;
RT "Cloning and sequencing of human gp330, a Ca(2+)-binding receptor
RT with potential intracellular signaling properties";
RL Eur. J. Biochem. 239:132-137 (1996).
RN [2]
RP SEQUENCE OF 2705-4453 FROM N.A.
RC TISSUE=Kidney;
RA Knaak C., Argraves W.S.;
RN Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 3833-4453 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95048397; PubMed=7959795;
RA Korenberg J.R., Argraves K.M., Chen X.N., Tran H.,
RA Strickland D.K., Argraves W.S.;
RT "Chromosomal localization of human genes for the LDL receptor family
RT member glycoprotein 330 (LRP2) and its associated protein RAP
RT (LRPAP1).";
RL Genomics 22:88-93 (1994).
RN [4]
RP SEQUENCE OF 4139-4406 FROM N.A.
RX MEDLINE=94244704; PubMed=8187828;
RA Lundgren S., Hjaelm G., Hellman P., Ek B., Juhlin C., Rastad J.,
RA Klarskog L., Akerstrom G., Rask L.;
RT "A protein involved in calcium sensing of the human parathyroid and
RT placental cytotrophoblast cells belongs to the LDL-receptor protein
RT superfamily";

RL	Exp. Cell Res. 212:344-350(1994).	FT	DOMAIN	65	105	LDL-RECEPTOR CLASS A 2.
RN	FUNCTION.	FT	DOMAIN	106	144	LDL-RECEPTOR CLASS A 3.
RA	MEDLINE=95286588; PubMed=7768901;	FT	DOMAIN	145	181	LDL-RECEPTOR CLASS A 4.
RX	Kounnas M.Z., Loukinova E.B., Stefansson S., Harmony J.A.K.,	FT	DOMAIN	182	219	LDL-RECEPTOR CLASS A 5.
RA	Brewer B.H., Strickland D.K., Agraves W.S.;	FT	DOMAIN	220	258	LDL-RECEPTOR CLASS A 6.
RT	"Identification of glycoprotein 330 as an endocytic receptor for	FT	DOMAIN	264	308	LDL-RECEPTOR CLASS A 7.
RT	apolipoprotein J/clusterin."	FT	DOMAIN	309	347	EGF-LIKE 1.
RL	J. Biol. Chem. 270:13070-13075(1995).	FT	DOMAIN	348	386	EGF-LIKE 2.
CC	-!- FUNCTION: BINDS SPECIFICALLY CLUSTERIN WITH HIGH AFFINITY. BUT	FT	DOMAIN	436	477	LDL-RECEPTOR CLASS B 1.
CC	ALSO LIGANDS IN COMMON WITH OTHER FAMILY MEMBERS: PLASMINOGEN	FT	DOMAIN	479	520	LDL-RECEPTOR CLASS B 2.
CC	EXTRACELLULAR MATRIX COMPONENTS, PLASMINOGEN ACTIVATOR-PLASMINOGEN	FT	DOMAIN	522	567	LDL-RECEPTOR CLASS B 3.
CC	ACTIVATOR INHIBITOR TYPE 1 COMPLEX, APOLOPROTEIN E-ENRICHED	FT	DOMAIN	569	612	LDL-RECEPTOR CLASS B 4.
CC	BETA-VLDL, LIPOPROTEIN LIPASE, LACTOFERRIN AND CALCIUM.	FT	DOMAIN	613	653	LDL-RECEPTOR CLASS B 5.
CC	-!- FUNCTION: RECEPTOR-MEDIATED UPTAKE OF POLYBASIC DRUGS SUCH AS	FT	DOMAIN	659	705	EGF-LIKE 3.
CC	APROTININ, AMINOGLYCOSIDES AND POLYMYXIN B (BY SIMILARITY).	FT	DOMAIN	753	794	LDL-RECEPTOR CLASS B 6.
CC	-!- FUNCTION: MAY PARTICIPATE IN REGULATION OF PARATHYROID-HORMONE AND	FT	DOMAIN	796	836	LDL-RECEPTOR CLASS B 7.
CC	PARA-THYROID-HORMONE-RELATED PROTEIN RELEASE.	FT	DOMAIN	838	880	LDL-RECEPTOR CLASS B 8.
CC	-!- SUBUNIT: Forms a multimeric complex together with a receptor-	FT	DOMAIN	882	924	LDL-RECEPTOR CLASS B 9.
CC	associated protein (RAP). Binds to ankyrin-repeat family A protein	FT	DOMAIN	970	1014	EGF-LIKE 4.
CC	2 (ANKRA2) (By similarity).	FT	DOMAIN	1024	1062	LDL-RECEPTOR CLASS A 8.
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.	FT	DOMAIN	1065	1103	LDL-RECEPTOR CLASS A 9.
CC	-!- TISSUE SPECIFICITY: ABSORPTIVE EPITHELIA, INCLUDING RENAL	FT	DOMAIN	1107	1145	LDL-RECEPTOR CLASS A 10.
CC	PROXIMAL TUBULES.	FT	DOMAIN	1147	1185	LDL-RECEPTOR CLASS A 11.
CC	-!- SIMILARITY: Contains 36 LDL-receptor class A domains.	FT	DOMAIN	1186	1224	LDL-RECEPTOR CLASS A 12.
CC	-!- SIMILARITY: Contains 37 LDL-receptor class B domains.	FT	DOMAIN	1228	1268	LDL-RECEPTOR CLASS A 13.
CC	-!- SIMILARITY: Contains 17 EGF-like domains.	FT	DOMAIN	1269	1307	LDL-RECEPTOR CLASS A 14.
CC	-----	FT	DOMAIN	1310	1350	LDL-RECEPTOR CLASS A 15.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	FT	DOMAIN	1349	1389	EGF-LIKE 5.
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	FT	DOMAIN	1390	1429	EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
CC	the European Bioinformatics Institute. There are no restrictions on its	FT	DOMAIN	1478	1519	LDL-RECEPTOR CLASS B 10.
CC	use by non-profit institutions as long as its content is in no way	FT	DOMAIN	1521	1562	LDL-RECEPTOR CLASS B 11.
CC	modified and this statement is not removed. Usage by and for commercial	FT	DOMAIN	1566	1608	LDL-RECEPTOR CLASS B 12.
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	FT	DOMAIN	1610	1653	LDL-RECEPTOR CLASS B 13.
CC	or send an email to license@isb-sib.ch).	FT	DOMAIN	1655	1695	LDL-RECEPTOR CLASS B 14.
CC	-----	FT	DOMAIN	1700	1741	EGF-LIKE 7.
DR	EMBL; U33837; AAB41649.1; -.	FT	DOMAIN	1790	1831	LDL-RECEPTOR CLASS B 15.
DR	EMBL; U04441; AAB02882.1; -.	FT	DOMAIN	1833	1881	LDL-RECEPTOR CLASS B 16.
DR	EMBL; S73145; AAB30825.1; -.	FT	DOMAIN	1883	1929	LDL-RECEPTOR CLASS B 17.
DR	PIR; I53413; I53413.	FT	DOMAIN	1931	1971	LDL-RECEPTOR CLASS B 18.
DR	HSSP; Q07954; ICR8.	FT	DOMAIN	1972	2012	LDL-RECEPTOR CLASS B 19.
DR	MIM; 600073; -.	FT	DOMAIN	2018	2059	EGF-LIKE 8.
DR	GO; GO:0005764; C:lysosome; TAS.	FT	DOMAIN	2107	2155	LDL-RECEPTOR CLASS B 20.
DR	GO; GO:0006629; P:lipid metabolism; TAS.	FT	DOMAIN	2157	2200	LDL-RECEPTOR CLASS B 21.
DR	GO; GO:0006486; P:protein amino acid glycosylation; TAS.	FT	DOMAIN	2202	2244	LDL-RECEPTOR CLASS B 22.
DR	GO; GO:0006898; P:receptor mediated endocytosis; TAS.	FT	DOMAIN	2246	2288	LDL-RECEPTOR CLASS B 23.
DR	InterPro; IPR000152; ASX_hydroxyl.	FT	DOMAIN	2290	2331	EGF-LIKE 9.
DR	InterPro; IPR001881; EGF_Ca.	FT	DOMAIN	2342	2383	LDL-RECEPTOR CLASS B 24.
DR	InterPro; IPR006209; EGF_like.	FT	DOMAIN	2431	2476	LDL-RECEPTOR CLASS B 25.
DR	InterPro; IPR002172; LDL_receptor_A.	FT	DOMAIN	2478	2517	LDL-RECEPTOR CLASS B 26.
DR	InterPro; IPR000033; LDL_receptor_rep.	FT	DOMAIN	2519	2561	LDL-RECEPTOR CLASS B 27.
DR	Pfam; PF00008; EGF; 13.	FT	DOMAIN	2563	2603	LDL-RECEPTOR CLASS B 28.
DR	Pfam; PF00057; ldl_recept_a; 36.	FT	DOMAIN	2604	2646	LDL-RECEPTOR CLASS B 29.
DR	Pfam; PF00058; ldl_recept_b; 37.	FT	DOMAIN	2651	2693	EGF-LIKE 10.
DR	SMART; SM00179; EGF_CA; 3.	FT	DOMAIN	2698	2738	LDL-RECEPTOR CLASS A 16.
DR	SMART; SM00192; LDLA; 36.	FT	DOMAIN	2739	2777	LDL-RECEPTOR CLASS A 17.
DR	SMART; SM00135; LV; 33.	FT	DOMAIN	2778	2819	LDL-RECEPTOR CLASS A 18.
DR	PROSITE; PS00010; ASX_HYDROXYL; 4.	FT	DOMAIN	2820	2861	LDL-RECEPTOR CLASS A 19.
DR	PROSITE; PS00022; EGF_1; 1.	FT	DOMAIN	2862	2901	LDL-RECEPTOR CLASS A 20.
DR	PROSITE; PS01186; EGF_2; 9.	FT	DOMAIN	2904	2945	LDL-RECEPTOR CLASS A 21.
DR	PROSITE; PS01187; EGF_CA; 3.	FT	DOMAIN	2946	2990	LDL-RECEPTOR CLASS A 22.
DR	PROSITE; PS01209; LDLRA_1; 31.	FT	DOMAIN	2991	3029	LDL-RECEPTOR CLASS A 23.
DR	PROSITE; PS00688; LDLRA_2; 36.	FT	DOMAIN	3030	3070	LDL-RECEPTOR CLASS A 24.
DR	Glycoprotein; Repeat; Endocytosis; Coated pits; Transmembrane;	FT	DOMAIN	3073	3110	LDL-RECEPTOR CLASS A 25.
KW	Receptor; EGF-like domain; Signal; SH3-binding; Polymorphism.	FT	DOMAIN	3111	3151	EGF-LIKE 11.
FT	SIGNAL	FT	DOMAIN	3152	3192	EGF-LIKE 12.
FT	CHAIN	FT	DOMAIN	3192	3230	LDL-RECEPTOR CLASS B 30.
FT	26 4655	FT	DOMAIN	3231	3280	LDL-RECEPTOR CLASS B 31.
FT	26 4655	FT	DOMAIN	3281	3331	LDL-RECEPTOR CLASS B 32.
FT	26 4655	FT	DOMAIN	3332	3375	LDL-RECEPTOR CLASS B 33.
FT	26 4655	FT	DOMAIN	3376	3418	LDL-RECEPTOR CLASS B 34.
FT	26 4655	FT	DOMAIN	3419	3459	EGF-LIKE 13.
FT	26 64	FT	DOMAIN	3460	3500	LDL-RECEPTOR CLASS A 26.
FT	26 64	FT	DOMAIN	3501	3551	LDL-RECEPTOR CLASS A 27.


```
FT DOMAIN 3592 3632 LDL-RECEPTOR CLASS A 28.
FT DOMAIN 3633 3673 LDL-RECEPTOR CLASS A 29.
FT DOMAIN 3676 3716 LDL-RECEPTOR CLASS A 30.
FT DOMAIN 3717 3756 LDL-RECEPTOR CLASS A 31.
FT DOMAIN 3757 3795 LDL-RECEPTOR CLASS A 32.
FT DOMAIN 3796 3834 LDL-RECEPTOR CLASS A 33.
FT DOMAIN 3840 3880 LDL-RECEPTOR CLASS A 34.
FT DOMAIN 3881 3922 LDL-RECEPTOR CLASS A 35.
FT DOMAIN 3926 3964 LDL-RECEPTOR CLASS A 36.

Query Match 17.8%; Score 86.5; DB 1; Length 4655;
Best Local Similarity 26.1%; Pred. No. 2.1;
Matches 23; Conservative 14; Mismatches 30; Indels 21; Gaps 5;

Qy 6 CGNEKYDSCGSKC-----DKKC-----KYDGVVEEDDEPNVPLVRVCHQD 49
Db 3745 CGDNDSEENACPRECTESEPVCVQCPISRWICDHYNDGDNDSR---DCENRTCHPE 3801

Qy 50 CV-CEEGFYRNKDKCVSAEDCELDNDM 76
Db 3802 YFQCTSGHCVHSELKCDGSADC-LDASD 3828

RESULT 11
ICE2_ASCSU STANDARD; PRT; 65 AA.
AC P07852;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Chymotrypsin/elastase isoforms 2 to 5.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoidea).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE.
RX MEDLINE=84255715; PubMed=6564898;
RA Babin D.R., Feanasky R.J., Goss S.M.;
RT "The isoforms of chymotrypsin/elastase from Ascaris lumbricoidea: the primary structure."
RL Arch. Biochem. Biophys. 232:143-161(1984).
CC -!- FUNCTION: Defend the organism against the host's proteinases.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 TIL (trypsin inhibitory-like) domain.
DR HSP; P07851; IEAI.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
KW Serine protease inhibitor.
FT DOMAIN 4 59 TIL.
FT DISULFID 4 37 BY SIMILARITY.
FT DISULFID 13 32 BY SIMILARITY.
FT DISULFID 16 28 BY SIMILARITY.
FT DISULFID 20 59 BY SIMILARITY.
FT DISULFID 39 53 BY SIMILARITY.
FT ACT SITE 30 31 REACTIVE BOND.
FT VARIANT 25 25 K -> N (IN INHIBITOR 2 AND 4).
FT VARIANT 40 40 T -> S (IN INHIBITOR 2 AND 4).
FT VARIANT 64 65 MISSING (IN INHIBITOR 2).
FT VARIANT 65 65 K -> E (IN INHIBITOR 3).
FT VARIANT 65 65 K -> R.
SQ SEQUENCE 65 AA; 7241 MW; B4E51CA166EA4BE3 CRC64;

Query Match 17.6%; Score 85.5; DB 1; Length 65;
Best Local Similarity 28.4%; Pred. No. 0.037;
Matches 19; Conservative 7; Mismatches 28; Indels 13; Gaps 3;

Qy 6 CGNEKYDSCGSKCCKKCKYDGVVEEDDEPNVPLVRVCHQDVCE--EGFYRNKDDK 63
Db 4 CGKNEVTECTG--CELKC-----GQDEKTPCALMCRPPSCCECTPGRMRTDGGK 52

Qy 64 CVSAEDC 70
||| :|
```

```
Db 53 CVPVSEC 59

RESULT 12
LDVR_CHK LDVR_CHK STANDARD; PRT; 863 AA.
AC P98165;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Very low-density lipoprotein receptor precursor (VLDL receptor)
DE (Vitellogenin receptor) (VTG receptor).
GN VLDLR OR VTGR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Ovary;
RX MEDLINE=95045409; PubMed=7957081;
RA Bujo H., Hermann M., Kaderli M.O., Jacobsen L., Sugawara S.,
RA Nimpf J., Yamamoto T., Schneider W.J.;
RT "Chicken oocyte growth is mediated by an eight ligand binding repeat member of the LDL receptor family."
RL EMBO J. 13:5165-5175(1994).
RN [2]
RP SEQUENCE OF 510-518; 546-554 AND 819-827.
RC STRAIN=White leghorn; TISSUE=Follicle membrane;
RX MEDLINE=92011638; PubMed=1655760;
RA Barber D.L., Sanders E.J., Aebersold R., Schneider W.J.;
RT "The receptor for yolk lipoprotein deposition in the chicken oocyte."
RL J. Biol. Chem. 266:18761-18770(1991).
CC -!- FUNCTION: Binds VLDL and transports it into cells by endocytosis.
CC In order to be internalized, the receptor-ligand complexes must first cluster into clathrin-coated pits. Binding to Reelin induces tyrosine phosphorylation of Dab1 and modulation of Tau phosphorylation (By similarity).
CC -!- SUBUNIT: Binds to the extracellular matrix protein Reelin (By similarity).
CC -!- TISSUE SPECIFICITY: ABUNDANT IN OOCYTES; MUCH LESS IN HEART AND SKELETAL MUSCLE.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 8 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X80207; CAA56505.1; -.
DR PIR; S51789; S51789.
DR HSP; P01130; IAJJ.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; LDL_receptor_rep.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00057; ldl_recept_a; 8.
DR Pfam; PF00058; ldl_recept_b; 5.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00192; LDLA; 8.
DR SMART; SM00135; LY; 5.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
```


QY 4 MCGGENEKYDCSGSKEC-DKKCKYDGVBEEDDEPNVPCLVRVCHQDCV-CEEGFYRNKD 61

CC In order to be internalized, the receptor-ligand complexes must
CC first cluster into clathrin-coated pits. Binding to Reelin induces
CC tyrosine phosphorylation of Dab1 and modulation of tau
CC phosphorylation (By similarity).
CC -|- SUBUNIT: Binds to the extracellular matrix protein Reelin (By
CC similarity).
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
CC -|- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P98155-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P98155-2; Sequence=VSP_004304;
CC -|- TISSUE SPECIFICITY: ABUNDANT IN HEART AND SKELETAL MUSCLE; ALSO
CC OVARY AND KIDNEY; NOT IN LIVER.
CC -|- SIMILARITY: Contains 8 LDL-receptor class A domains.
CC -|- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC -|- SIMILARITY: Contains 3 EGF-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L20470; AAA53684.1; -;
DR EMBL; D16532; BAA03969.1; JOINED.
DR EMBL; D16495; BAA03969.1; JOINED.
DR EMBL; D16508; BAA03969.1; JOINED.
DR EMBL; D16510; BAA03969.1; JOINED.
DR EMBL; D16514; BAA03969.1; JOINED.
DR EMBL; D16516; BAA03969.1; JOINED.
DR EMBL; D16518; BAA03969.1; JOINED.
DR EMBL; D16520; BAA03969.1; JOINED.
DR EMBL; D16522; BAA03969.1; JOINED.
DR EMBL; D16523; BAA03969.1; JOINED.
DR EMBL; D16524; BAA03969.1; JOINED.
DR EMBL; D16525; BAA03969.1; JOINED.
DR EMBL; D16526; BAA03969.1; JOINED.
DR EMBL; D16527; BAA03969.1; JOINED.
DR EMBL; D16528; BAA03969.1; JOINED.
DR EMBL; D16529; BAA03969.1; JOINED.
DR EMBL; D16530; BAA03969.1; JOINED.
DR EMBL; D16531; BAA03969.1; JOINED.
DR EMBL; S73849; AAB31735.1; -;
DR EMBL; D16493; BAA03945.1; -;
DR EMBL; D16494; BAA03946.1; -;
DR EMBL; L22431; AAB61344.1; -;
DR F1R; A49729; A49729.
DR HSGP; P01130; IAJJ.
DR Genew; HGNC:12698; VLDLR.
DR MIM; 192977; -;
DR GO; GO:0005041; F:low-density lipoprotein receptor activity; TAS.
DR GO; GO:0007633; P:memory; TAS.
DR GO; GO:0007399; P:neurogenesis; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR00152; Asx hydroxyl.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR006309; EGF-like.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00057; ldl_recept_a; 8.
DR Pfam; PF00058; ldl_recept_b; 5.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00192; LDLA; 8.
DR SMART; SM00135; LY; 5.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 3.

DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS01209; LDLA_1; 8.
DR PROSITE; PS00068; LDLA_2; 8.
KW Glycoprotein; VLDL; Cholesterol metabolism; Lipid transport;
KW Endocytosis; Coated pits; Transmembrane; Receptor; Signal;
KW EGF-like domain; Repeat; Alternative splicing; Polymorphism.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 873 VERY LOW-DENSITY LIPOPROTEIN RECEPTOR.
FT DOMAIN 28 797 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 798 819 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 820 873 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 31 69 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 70 110 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 111 151 LDL-RECEPTOR CLASS A 4.
FT DOMAIN 152 190 LDL-RECEPTOR CLASS A 5.
FT DOMAIN 191 231 LDL-RECEPTOR CLASS A 6.
FT DOMAIN 237 275 LDL-RECEPTOR CLASS A 7.
FT DOMAIN 276 314 LDL-RECEPTOR CLASS A 8.
FT DOMAIN 316 355 EGF-LIKE 1.
FT DOMAIN 356 395 EGF-LIKE 2.
FT DOMAIN 396 435 EGF-LIKE 3.
FT REPEAT 439 480 EGF-LIKE 1. CALCIUM-BINDING (POTENTIAL).
FT REPEAT 481 524 LDL-RECEPTOR CLASS B 1.
FT REPEAT 525 567 LDL-RECEPTOR CLASS B 2.
FT REPEAT 568 611 LDL-RECEPTOR CLASS B 3.
FT REPEAT 612 654 LDL-RECEPTOR CLASS B 4.
FT REPEAT 655 696 LDL-RECEPTOR CLASS B 5.
FT DOMAIN 702 750 EGF-LIKE 3.
FT DOMAIN 751 790 CLUSTERED O-LINKED OLIGOSACCHARIDES
FT SITE 832 837 (BY SIMILARITY).
FT DISULFID 33 45 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 40 58 BY SIMILARITY.
FT DISULFID 52 67 BY SIMILARITY.
FT DISULFID 72 84 BY SIMILARITY.
FT DISULFID 79 97 BY SIMILARITY.
FT DISULFID 91 108 BY SIMILARITY.
FT DISULFID 113 127 BY SIMILARITY.
FT DISULFID 120 140 BY SIMILARITY.
FT DISULFID 134 149 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
FT DISULFID 161 179 BY SIMILARITY.
FT DISULFID 173 188 BY SIMILARITY.
FT DISULFID 193 205 BY SIMILARITY.
FT DISULFID 200 218 BY SIMILARITY.
FT DISULFID 212 229 BY SIMILARITY.
FT DISULFID 239 251 BY SIMILARITY.
FT DISULFID 258 264 BY SIMILARITY.
FT DISULFID 264 273 BY SIMILARITY.
FT DISULFID 278 290 BY SIMILARITY.
FT DISULFID 285 303 BY SIMILARITY.
FT DISULFID 297 312 BY SIMILARITY.
FT DISULFID 318 331 BY SIMILARITY.
FT DISULFID 326 344 BY SIMILARITY.
FT DISULFID 338 355 BY SIMILARITY.
FT DISULFID 360 371 BY SIMILARITY.
FT DISULFID 367 380 BY SIMILARITY.
FT DISULFID 382 394 BY SIMILARITY.
FT DISULFID 400 410 BY SIMILARITY.
FT DISULFID 406 419 BY SIMILARITY.
FT DISULFID 421 434 BY SIMILARITY.
FT DISULFID 706 719 BY SIMILARITY.
FT DISULFID 715 734 BY SIMILARITY.

Query Match 17.0%; Score 82.5; DB 1; Length 873;
Best Local Similarity 31.7%; Pred. No. 0.96;
Matches 26; Conservative 7; Mismatches 38; Indels 11; Gaps 4;
Qy 5 QCGENEKYS-----CGSKC-DKCKYDGVDEEDDEPNVPCIVRVCHQD-CVCEE 54
Dy 228 QCGRPVHTKCPASEIQCGSGCIHKWKCDGDPCKGSDVNCPSRTCRPDQFECED 287
Qy 55 GFYRNKDDKCVSAEDCCLDND 76


```

FT DISULFID 297 312 BY SIMILARITY.
FT DISULFID 318 331 BY SIMILARITY.
FT DISULFID 326 344 BY SIMILARITY.
FT DISULFID 338 355 BY SIMILARITY.
FT DISULFID 360 371 BY SIMILARITY.
FT DISULFID 367 380 BY SIMILARITY.
FT DISULFID 382 394 BY SIMILARITY.
FT DISULFID 400 410 BY SIMILARITY.
FT DISULFID 406 419 BY SIMILARITY.
FT DISULFID 421 434 BY SIMILARITY.
FT DISULFID 706 719 BY SIMILARITY.
FT DISULFID 715 734 BY SIMILARITY.
FT DISULFID 736 749 BY SIMILARITY.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 765 765 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 781 781 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 161 161 C -> G (IN REF. 1).
FT CONFLICT 262 262 P -> L (IN REF. 3).
FT CONFLICT 297 297 C -> S (IN REF. 2).
SQ SEQUENCE 873 AA; 96372 MW; 08F09F93825195CB CRC64;

Query Match 17.0%; Score 82.5; DB 1; Length 873;
Best Local Similarity 31.7%; Pred. No. 0.96;
Matches 26; Conservative 7; Mismatches 38; Indels 11; Gaps 4;

Qy 5 QCGENEKYDS-----CGSKEC-DKKCKYGVVEEDDEPNVPCLVRYCHOD-CVCEE 54
Dy 228 QCGQPVHHTKCTSETSCGSGECHKWRCGDGDKGSDGVNCPSTRCPDQFCEC 287
Qy 55 GFYRNKDDKCVSABDCELDND 76
Dy 288 GSCIHSRQCNGIRDG-VGSD 308

RESULT 15
ID LVDR RABIT STANDARD; PRT; 873 AA.
AC P35953;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Very low-density lipoprotein receptor precursor (VLDL receptor).
GN VLDLR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9302842; PubMed=1384047;
RA Takahashi S., Kawarabayashi Y., Nakai T., Sakai J., Yamamoto T.;
RT "Rabbit very low density lipoprotein receptor: a low density
RT lipoprotein receptor-like protein with distinct ligand specificity.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:9252-9256(1992).
CC -!- FUNCTION: Binds VLDL and transports it into cells by endocytosis.
CC In order to be internalized, the receptor-ligand complexes must
CC first cluster into clathrin-coated pits. Binding to Reelin induces
CC tyrosine phosphorylation of Dab1 and modulation of tau
CC phosphorylation (By similarity).
CC -!- SUBUNIT: Binds to the extracellular matrix protein Reelin (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: ABUNDANT IN HEART, MUSCLE, AND ADIPOSE
CC TISSUE.
CC -!- SIMILARITY: Contains 8 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 6 LDL-receptor class B domains.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC

entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; D11100; BAA01874.1; -.
CC PIR; A46286; QRRBYD.
CC HSSP; P01130; 1AJJ.
CC InterPro; IPR000152; Asx hydroxyl.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR002172; LDL_receptor_A.
CC InterPro; IPR00033; LDL_receptor_rep.
CC Pfam; PF00057; ldl_recept_a; 8.
CC Pfam; PF00058; ldl_recept_b; 5.
CC PRINTS; PR00261; LDLRECEPTOR.
CC SMART; SM00179; EGF_CA; 1.
CC SMART; SM00192; LDLA; 8.
CC SMART; SM00135; LY; 5.
CC PROSITE; PS00010; ASX_HYDROXYL; 2.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 3.
CC PROSITE; PS01187; EGF_CA; 1.
CC PROSITE; PS01209; LDLRA_1; 8.
CC PROSITE; PS00068; LDLRA_2; 8.
CC Glycoprotein; VLDL; Cholesterol metabolism; Lipid transport;
CC Endocytosis; Coated pits; Transmembrane; Receptor; Signal;
CC EGF-like domain; Repeat.
CC SIGNAL 1 27 POTENTIAL.
CC CHAIN 28 873 VERY LOW-DENSITY LIPOPROTEIN RECEPTOR.
CC DOMAIN 28 797 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 798 819 POTENTIAL.
CC DOMAIN 820 873 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 31 69 LDL-RECEPTOR CLASS A 1.
CC DOMAIN 70 110 LDL-RECEPTOR CLASS A 2.
CC DOMAIN 111 151 LDL-RECEPTOR CLASS A 3.
CC DOMAIN 152 190 LDL-RECEPTOR CLASS A 4.
CC DOMAIN 191 231 LDL-RECEPTOR CLASS A 5.
CC DOMAIN 237 275 LDL-RECEPTOR CLASS A 6.
CC DOMAIN 276 314 LDL-RECEPTOR CLASS A 7.
CC DOMAIN 316 355 LDL-RECEPTOR CLASS A 8.
CC DOMAIN 356 395 EGF-LIKE 1.
CC DOMAIN 396 435 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
CC REPEAT 439 480 LDL-RECEPTOR CLASS B 1.
CC REPEAT 481 524 LDL-RECEPTOR CLASS B 2.
CC REPEAT 525 567 LDL-RECEPTOR CLASS B 3.
CC REPEAT 568 611 LDL-RECEPTOR CLASS B 4.
CC REPEAT 612 654 LDL-RECEPTOR CLASS B 5.
CC REPEAT 655 696 LDL-RECEPTOR CLASS B 6.
CC DOMAIN 702 750 EGF-LIKE 3.
CC DOMAIN 751 790 CLUSTERED O-LINKED OLIGOSACCHARIDES.
CC SITE 832 837 ENDOCYTOSIS SIGNAL (POTENTIAL).
CC DISULFID 33 45 BY SIMILARITY.
CC DISULFID 40 58 BY SIMILARITY.
CC DISULFID 52 67 BY SIMILARITY.
CC DISULFID 72 84 BY SIMILARITY.
CC DISULFID 79 97 BY SIMILARITY.
CC DISULFID 91 108 BY SIMILARITY.
CC DISULFID 113 127 BY SIMILARITY.
CC DISULFID 120 140 BY SIMILARITY.
CC DISULFID 134 149 BY SIMILARITY.
CC DISULFID 154 166 BY SIMILARITY.
CC DISULFID 161 179 BY SIMILARITY.
CC DISULFID 173 188 BY SIMILARITY.
CC DISULFID 193 205 BY SIMILARITY.
CC DISULFID 200 218 BY SIMILARITY.
CC DISULFID 212 229 BY SIMILARITY.
CC DISULFID 239 251 BY SIMILARITY.
CC DISULFID 246 264 BY SIMILARITY.
CC DISULFID 258 273 BY SIMILARITY.
CC DISULFID 278 290 BY SIMILARITY.
CC DISULFID 285 303 BY SIMILARITY.
CC DISULFID 297 312 BY SIMILARITY.
CC DISULFID 318 331 BY SIMILARITY.
```

Search completed: December 8, 2003, 09:14:12
Job time : 24.0606 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 88.2424 Seconds
(without alignments)
245.646 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1 KATMQCENEKYDCSGKCE.....VSAEDCELDNMDFIYFGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rhodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	91	5 Q16938	Q16938 ancylostoma
2	196.5	40.4	102	5 Q962V8	Q962V8 ancylostoma
3	158.5	32.6	76	5 Q909B1	Q909B1 ancylostoma
4	139.5	28.7	98	5 Q16939	Q16939 ancylostoma
5	137.5	28.3	77	5 Q16935	Q16935 ancylostoma
6	137.5	28.3	91	5 Q16947	Q16947 ancylostoma
7	137.5	28.3	100	5 Q16940	Q16940 ancylostoma
8	125	25.7	432	5 Q814B8	Q814B8 caenorhabdi
9	125	25.7	537	5 Q901T6	Q901T6 caenorhabdi
10	124	25.5	102	12 Q9DQD5	Q9DQD5 microplitis
11	123.5	25.4	84	5 Q816Q5	Q816Q5 anopheles s
12	120.5	24.8	499	11 Q88714	Q88714 mus musculu
13	119.5	24.6	154	5 Q81TP8	Q81TP8 oesophagost
14	117	24.1	77	5 Q8TOW0	Q8TOW0 pimpha hypo
15	113	23.3	222	5 Q16620	Q16620 caenorhabdi
16	112	23.0	249	5 Q45764	Q45764 caenorhabdi

17	111.5	22.9	69	5 Q8MX86	Q8MX86 glossina mo
18	111	22.8	313	11 Q8K3U2	Q8K3U2 mus musculu
19	110	22.6	796	5 Q9U1T5	Q9U1T5 caenorhabdi
20	109.5	22.5	1642	5 Q62055	Q62055 caenorhabdi
21	106	21.8	731	5 Q814B9	Q814B9 caenorhabdi
22	105.5	21.7	145	5 Q18156	Q18156 caenorhabdi
23	100	20.6	5374	11 Q99ND0	Q99ND0 mus musculu
24	99	20.4	1036	5 Q97378	Q97378 strongyloce
25	98.5	20.3	86	5 Q817M1	Q817M1 caenorhabdi
26	98.5	20.3	135	5 Q18158	Q18158 caenorhabdi
27	98	20.2	137	5 Q18157	Q18157 caenorhabdi
28	97	20.0	3843	5 Q9VU94	Q9VU94 drosophila
29	97	20.0	3843	5 Q9U5D0	Q9U5D0 drosophila
30	96.5	19.9	869	13 Q42126	Q42126 xenopus lae
31	96	19.8	226	12 Q98332	Q98332 microplitis
32	96	19.8	315	12 Q98331	Q98331 microplitis
33	95.5	19.7	166	5 Q16488	Q16488 caenorhabdi
34	94.5	19.4	98	5 Q94162	Q94162 caenorhabdi
35	93.5	19.2	2108	13 Q98U19	Q98U19 gallus gall
36	92.5	19.0	2551	4 Q8WQ8	Q8WQ8 homo sapien
37	92.5	19.0	2551	4 Q8TUG9	Q8TUG9 homo sapien
38	91.5	18.8	62	5 Q77419	Q77419 ascaris suu
39	90.5	18.6	195	5 Q18805	Q18805 caenorhabdi
40	89.5	18.4	2559	11 Q8R4U0	Q8R4U0 mus musculu
41	87.5	18.0	949	5 P90956	P90956 caenorhabdi
42	87	17.9	92	5 Q9GPC4	Q9GPC4 drosophila
43	86.5	17.8	807	5 Q8MP01	Q8MP01 halocynthia
44	86	17.7	457	5 Q9NC91	Q9NC91 strongyloce
45	85	17.5	461	11 P97883	P97883 rattus norv

ALIGNMENTS

RESULT 1

Q16938	PRELIMINARY;	PRT;	91 AA.
ID Q16938			
AC Q16938;			
DC 01-NOV-1996 (TREMBLrel. 01, Created)			
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE Anti-coagulant protein C2 precursor (Fragment)			
OS Ancylostoma caninum (Dog hookworm)			
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;			
OC Ancylostomatidae; Ancylostomatinae; Ancylostominae; Ancylostoma.			
OX NCBI_TaxID=29170;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=98298519; PubMed=9634780;			
RA Jespers L.S., Messens J.H., De Keyser A., Beckhout D.,			
RA van den Brande I., Gansemans Y.G., Lauwereys M.J., Vlasuk G.P.,			
RA Stanssens P.E.;			
RT "surface expression and ligand-based selection of cDNAs fused to			
RT filamentous phage gene VI."			
RL Biotechnology 13:378-382(1995).			
RN [2]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=96312555; PubMed=8700900;			
RA Stanssens P., Bergum P.W., Gansemans Y., Jaspers L., Laroche Y.,			
RA Huang S., Maki S., Messens J., Lauwereys M., Cappello M., Hotez P.J.,			
RA Lasters I., Vlasuk G.P.;			
RT "Anticoagulant repertoire of the hookworm Ancylostoma caninum.";			
RL Proc. Natl. Acad. Sci. U.S.A. 93:2149-2154(1996).			
DR EMBL; U30793; AAC47080.1; -			
DR HSSP; P56682; ICCV.			
DR InterPro; IPR006209; EGF like.			
DR InterPro; IPR002919; TIL_Cysrich.			
DR Pfam; PF01826; TIL; 1.			
DR PROSITE; PS01186; EGF_2; 1.			
KW Signal.			
FT NON_TER	1	1	
FT SIGNAL	<1	7	POTENTIAL.
FT CHAIN	8	91	ANTI-COAGULANT PROTEIN C2.

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anticoagulant peptide (Fragment).
OS Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93274118; PubMed=8501344;
RA Cappello M., Clyne L.P., McPhedran P., Hotez P.J.;
RT "Ancylostoma factor Xa inhibitor: partial purification and its
RT identification as a major hookworm-derived anticoagulant in vitro";
RL J. Infect. Dis. 167:1474-1477(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97039675; PubMed=8885227;
RA Cappello M., Hawdon J.M., Jones B.F., Kennedy W.P., Hotez P.J.;
RT "Ancylostoma caninum anticoagulant peptide: cloning by PCR and
RT expression of soluble, active protein in E. coli.";
RL Mol. Biochem. Parasitol. 80:113-117(1996).
DR EMBL; U18305; AAC47318.1; -
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
DR PROSITE; PS01186; EGF_2; 1.
FT NON_TER 1
SQ SEQUENCE 77 AA; 8746 MW; 7F9294472EEB1119 CRC64;
Query Match 28.3%; Score 137.5; DB 5; Length 77;
Best Local Similarity 42.5%; Pred. No. 9.2e-09;
Matches 31; Conservative 11; Mismatches 26; Indels 5; Gaps 3;
QY 1 KATWCCGENEYKDSGCS-KECDKKCKYDGVVEEDDEEPNVCILRVCHQDCVCEEGFYRN 59
DB 1 KAYPECGENWLDGCTQKPCAEKCNPEEPDEPFCRSRGCLLPPA---CVCKDGFYRD 57
QY 60 K-DDKCVSAEDCE 71
DB 58 TVIGDCVREEDC 70
RESULT 6
ID Q16947 PRELIMINARY; PRT; 91 AA.
AC Q16947;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anticoagulant peptide isoform 2.
GN ACAP-2
OS Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95320231; PubMed=7597095;
RA Cappello M., Vlasuk G.P., Bergum P.W., Huang S., Hotez P.J.;
RT "Ancylostoma caninum anticoagulant peptide: a hookworm-derived
RT inhibitor of human coagulation factor Xa";
RL Proc. Natl. Acad. Sci. U.S.A. 92:6152-6156(1995).
DR EMBL; U46550; AAB87584.1; -
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.

DR PROSITE; PS01186; EGF_2; 1.
SQ SEQUENCE 91 AA; 10324 MW; E627F11E89B3EB4A CRC64;
Query Match 28.3%; Score 137.5; DB 5; Length 91;
Best Local Similarity 42.5%; Pred. No. 1.1e-08;
Matches 31; Conservative 11; Mismatches 26; Indels 5; Gaps 3;
QY 1 KATWCCGENEYKDSGCS-KECDKKCKYDGVVEEDDEEPNVCILRVCHQDCVCEEGFYRN 59
DB 15 KAYPECGENWLDGCTQKPCAEKCNPEEPDEPFCRSRGCLLPPA---CVCKDGFYRD 71
QY 60 K-DDKCVSAEDCE 71
DB 72 TVIGDCVREEDC 84
RESULT 7
ID Q16940 PRELIMINARY; PRT; 100 AA.
AC Q16940;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anti-coagulant protein 5 precursor.
OS Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312555; PubMed=8700900;
RA Stanssens P., Bergum P.W., Ganssems Y., Jaspers L., Laroche Y.,
RA Huang S., Maki S., Messens J., Lauwereys M., Cappello M., Hotez P.J.,
RA Lasters I., Vlasuk G.P.;
RT "Anticoagulant repertoire of the hookworm Ancylostoma caninum";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2149-2154(1996).
DR EMBL; U30795; AAC47082.1; -
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
DR PROSITE; PS01186; EGF_2; 1.
KW Signal.
FT CHAIN 1 23 POTENTIAL.
FT CHAIN 23 100 ANTICOAGULANT PROTEIN 5.
SQ SEQUENCE 100 AA; 11359 MW; DE9CB76729719A42 CRC64;
Query Match 28.3%; Score 137.5; DB 5; Length 100;
Best Local Similarity 42.5%; Pred. No. 1.2e-08;
Matches 31; Conservative 11; Mismatches 26; Indels 5; Gaps 3;
QY 1 KATWCCGENEYKDSGCS-KECDKKCKYDGVVEEDDEEPNVCILRVCHQDCVCEEGFYRN 59
DB 24 KAYPECGENWLDGCTQKPCAEKCNPEEPDEPFCRSRGCLLPPA---CVCKDGFYRD 80
QY 60 K-DDKCVSAEDCE 71
DB 81 TVIGDCVREEDC 93
RESULT 8
ID Q814B8 PRELIMINARY; PRT; 432 AA.
AC Q814B8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein Y69H2.3d.
GN Y69H2.3d.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]

DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE GASTRIC MUCIN-LIKE protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stomach;
RA Tomasetto C., Masson R., Wendling C., Lefebvre O., Chenard M.P.,
RA Rio M.C.;
RT "Identification of interactions between trefoil peptides and members
RT of the mucin protein family using the yeast two-hybrid system.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ010752; CAA09343.1; -
DR HSSP; P56682; ICCV.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF01826; TIL; 1.
DR Pfam; PF00094; vwd; 1.
DR SMART; SM00216; VMD; 1.
FT NON_TER 1
FT NON_TER 499
SQ SEQUENCE 499 AA; 54190 MW; 04F89EF4F23EE61E CRC64;

Query Match 24.8%; Score 120.5; DB 11; Length 499;
Best Local Similarity 30.1%; Pred. No. 5.7e-06;
Matches 25; Conservative 8; Mismatches 35; Indels 15; Gaps 2;

QY 6 CGENEKDYSCGSKCECDK-----KCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFGYR 58
DB 14 CPEPKTFQSCQSSEDEKFGACAPTQMLAT-----GIDCVPTKESGGCVCPKGLYE 65

QY 59 NKDDKCVSAEDCELDNMDFIYPG 81
DB 66 NSDQGVPAEPCDYGVSYPG 88

RESULT 13
Q81TP8 PRELIMINARY; PRT; 154 AA.
ID Q81TP8;
AC Q81TP8;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Putative trypsin-like inhibitor protein precursor.
GN MCRP.
OS Oesophagostomum dentatum.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Strongyloidea; Chabertiidae; Oesophagostomum.
OX NCBI_TaxID=61180;
RN [1]
RP SEQUENCE FROM N.A.
RA Boag P.R., Ranganathan S., Newton S.E., Gasser R.B.;
RT "Identification of a male-specific nematode protein with two trypsin
RT like inhibitor domains.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF399936; AAN32637.1; -
KW SIGNAL.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 154 PUTATIVE TRYPSIN-LIKE INHIBITOR PROTEIN.
SQ SEQUENCE 154 AA; 16564 MW; A7F566E3957DA819 CRC64;

Query Match 24.6%; Score 119.5; DB 5; Length 154;
Best Local Similarity 36.9%; Pred. No. 2.4e-06;
Matches 24; Conservative 8; Mismatches 20; Indels 13; Gaps 3;

QY 6 CGENEKDYSCGSKCECDKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFGYRNDKCV 65
DB 26 CGENEYINPCGN-HCEDTCSF-----TRRGCIAMCGPAACVCKEGFYRNSAGKC- 73

QY 66 SAEDC 70
DB 74 -TKDC 77

RESULT 14

Q8TOW0 PRELIMINARY; PRT; 77 AA.
ID Q8TOW0;
AC Q8TOW0;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Cysteine-rich venom protein 6 precursor.
GN CVP6.
OS Pimpla hypochondriaca.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Ichneumonoidea;
OC Ichneumonidae; Pimplinae; Pimpla.
OX NCBI_TaxID=135724;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom secretory;
RA Parkinson N.M., Richards E.H., Conyers C., Smith I., Edwards J.P.;
RT "Analysis of venom constituents from the parasitoid wasp Pimpla
RT hypochondriaca and cloning of a cDNA encoding a venom protein.";
RL Insect Biochem. Mol. Biol. 0:0-0(0).
DR EMBL; AJ438997; CAD27742.1; -
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
KW SIGNAL.
FT SIGNAL 1 21
FT CHAIN 22 77 CYSTEINE-RICH VENOM PROTEIN 6.
SQ SEQUENCE 77 AA; 8184 MW; 689CB5F40D63DD6 CRC64;

Query Match 24.1%; Score 117; DB 5; Length 77;
Best Local Similarity 35.4%; Pred. No. 2.3e-06;
Matches 23; Conservative 8; Mismatches 22; Indels 12; Gaps 3;

QY 6 CGENEKDYSCGSKCECDKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFGYRNDKCV 65
DB 24 CGPNRVYKSGT-GCPETC-----ENPDPC-DRACHQGCFCCKGLQIDGNCI 71

QY 66 SAEDC 70
DB 72 SPDKC 76

RESULT 15

O16620 PRELIMINARY; PRT; 222 AA.
ID O16620
AC O16620;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein F36H9.4.
GN F36H9.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R., Dante M., Kramer J., Twyman B.;

RT "The sequence of *C. elegans* cosmid F36H9.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA Waterston R.;

Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF016668; AAO12440.1; -.

KW Hypothetical protein.

AN hypochlorite protein.
SQ SEQUENCE 222 AA; 25752 MW; A9A999F0D2EEA03F CRC64;

Query Match

23.38: Score 113: DB 5: Length 222:

Query Match: 23.3%, Score: 115, Length: 222,
Best Local Similarity: 34.7%, Pred. No. 2e-05;

[illegible]

57 YRNKDDKCVSAE 68 QY

[illegible]

Search completed: December 8, 2003, 09:18:37
Job time : 90.2424 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:18:46 ; Search time 74.6667 seconds
(without alignments)
209.232 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1-KATMOGGENEYKDCSKEC.....VSAEDCELDNMDFIYPGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	84	11	US-09-498-272-59
2	486	100.0	91	11	US-09-498-272-128
3	212	43.6	82	11	US-09-498-272-25
4	212	43.6	82	11	US-09-498-272-49
5	212	43.6	171	11	US-09-498-272-62
6	202	41.6	86	11	US-09-498-272-46
7	201	41.4	89	11	US-09-498-272-48
8	201	41.4	108	11	US-09-498-272-24
9	199	40.9	86	11	US-09-498-272-45
10	197.5	40.6	88	11	US-09-498-272-43
11	195.5	40.2	84	11	US-09-498-272-50
12	195.5	40.2	162	11	US-09-498-272-63
13	193.5	39.8	84	11	US-09-498-272-51
14	193.5	39.8	162	11	US-09-498-272-64
15	186	38.3	83	11	US-09-498-272-52

Sequence 27, Appl
Sequence 65, Appl
Sequence 44, Appl
Sequence 47, Appl
Sequence 58, Appl
Sequence 23, Appl
Sequence 6, Appl
Sequence 41, Appl
Sequence 8, Appl
Sequence 21, Appl
Sequence 4, Appl
Sequence 40, Appl
Sequence 28, Appl
Sequence 56, Appl
Sequence 7, Appl
Sequence 20, Appl
Sequence 57, Appl
Sequence 22, Appl
Sequence 53, Appl
Sequence 26, Appl
Sequence 55, Appl
Sequence 54, Appl
Sequence 40, Appl
Sequence 42, Appl
Sequence 1, Appl
Sequence 2, Appl
Sequence 3, Appl
Sequence 4, Appl
Sequence 60, Appl

US-09-498-272-27
US-09-498-272-65
US-09-498-272-44
US-09-498-272-47
US-09-498-272-58
US-09-498-272-23
US-09-498-272-6
US-09-498-272-41
US-09-498-272-8
US-09-498-272-21
US-09-498-272-4
US-09-498-272-40
US-09-498-272-28
US-09-498-272-56
US-09-498-272-7
US-09-498-272-20
US-09-498-272-57
US-09-498-272-22
US-09-498-272-53
US-09-498-272-26
US-09-498-272-55
US-09-498-272-54
US-10-087-887-40
US-09-498-272-42
US-10-174-151-1
US-10-174-151-2
US-10-174-151-3
US-10-174-151-4
US-10-114-774-4
US-09-498-272-60

16 186 38.3 102 11
17 186 38.3 161 11
18 184.5 38.0 87 11
19 36.6 78 11
20 159 32.7 77 11
21 159 32.7 77 11
22 139.5 28.7 96 11
23 139.5 28.7 75 11
24 139.5 28.7 79 11
25 139.5 28.7 98 11
26 137.5 28.3 77 11
27 137.5 28.3 77 11
28 137.5 28.3 78 11
29 137.5 28.3 78 11
30 137.5 28.3 81 11
31 137.5 28.3 100 11
32 137 28.2 75 11
33 137 28.2 94 11
34 136.5 28.1 78 11
35 136 28.0 75 11
36 134.5 27.7 77 11
37 134.5 27.7 78 11
38 118 24.3 67 12
39 111 22.8 74 11
40 100 20.6 69 12
41 100 20.6 69 12
42 100 20.6 71 12
43 100 20.6 73 12
44 87 17.9 115 12
45 84.5 17.4 58 11

ALIGNMENTS

RESULT 1

US-09-498-272-59
; Sequence 59, Application US/09498272
; Publication No. US20030113890A1

GENERAL INFORMATION:

APPLICANT: Viasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Ganssema, Yannick Georges Jozef
Moyle, Matthew
Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS: .

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/498,272

FILING DATE: 04-Feb-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231

FILING DATE: October 17, 1995

```
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-09-498-272-59

Query Match 100.0%; Score 486; DB 11; Length 84;
Best Local Similarity 100.0%; Pred. No. 8.6e-42;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCCEGFYRNK 60
Db 1 KATMQCGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCCEGFYRNK 60

Qy 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 2
US-09-498-272-128
; Sequence 128, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
```

```
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 128:
US-09-498-272-128

Query Match 100.0%; Score 486; DB 11; Length 91;
Best Local Similarity 100.0%; Pred. No. 9.4e-42;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCCEGFYRNK 60
Db 8 KATMQCGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCCEGFYRNK 67

Qy 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 68 DDKCVSAEDCELDNMDFIYPGTRN 91

RESULT 3
US-09-498-272-25
; Sequence 25, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
```

/ /
/ COUNTRY: U.S.A.
/ ZIP: 90071
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: IBM P.C. DOS 5.0
/ SOFTWARE: Word Perfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/498,272
/ FILING DATE: 04-Feb-2000
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/13231
/ FILING DATE: October 17, 1995
/ APPLICATION NUMBER: 08/486,399
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/486,397
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/465,380
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/461,965
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/326,110
/ FILING DATE: October 18, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BIGGS, SUZANNE L.
/ REGISTRATION NUMBER: 30,158
/ REFERENCE/DOCKET NUMBER: 216/270
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 67-3510
/ INFORMATION FOR SEQ ID NO: 25:
/ SEQUENCE CHARACTERISTICS:
/ TYPE: amino acid
/ LENGTH: 82 amino acids
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: Ancylostoma ceylanicum
/ SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-498-272-25

Query Match 43.6%; Score 212; DB 11; Length 82;
Best Local Similarity 50.0%; Pred. No. 3e-14;
Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;
QY 6 CGENEKYDCGS-KECDKKYGVDEEDDEPNVPLVRVCHQD--CVCBEGFYRNKDD 62
Db 4 CGSNERYSDCGNDKQCKERKCNEDDYKGD-----ACRSHVCERPACVCEGDFYRNKKG 58

QY 63 KCVSAEDCELDNMDFI 78
Db 59 SCVESDDCEYDNMDFI 74

RESULT 4

US-09-498-272-49
/ Sequence 49, Application US/09498272
/ Publication No. US20030113890A1
/ GENERAL INFORMATION:
/ APPLICANT: Vlasuk, George Phillip
/ Stanssens, Patrick Eric Hugo
/ Messens, Joris Hilda Lieven
/ Laroche, Yves Rene
/ Lauwereys, Marc Josef
/ Jespers, Laurent Stephane
/ Ganssens, Yannick Georges Josef
/ Moyle, Matthew
/ Bergum, Peter W.
/ TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
/ INHIBITORS AND ANTICOAGULANT
/ PROTEIN

/ /
/ NUMBER OF SEQUENCES: 356
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Lyon & Lyon
/ STREET: 633 West Fifth Street
/ Suite 4700
/ CITY: Los Angeles
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 90071
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: IBM P.C. DOS 5.0
/ SOFTWARE: Word Perfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/498,272
/ FILING DATE: 04-Feb-2000
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/13231
/ FILING DATE: October 17, 1995
/ APPLICATION NUMBER: 08/486,399
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/486,397
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/465,380
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/461,965
/ FILING DATE: June 5, 1995
/ APPLICATION NUMBER: 08/326,110
/ FILING DATE: October 18, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BIGGS, SUZANNE L.
/ REGISTRATION NUMBER: 30,158
/ REFERENCE/DOCKET NUMBER: 216/270
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (213) 489-1600
/ TELEFAX: (213) 955-0440
/ TELEX: 67-3510
/ INFORMATION FOR SEQ ID NO: 49:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 82 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: Ancylostoma ceylanicum
/ SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-09-498-272-49

Query Match 43.6%; Score 212; DB 11; Length 82;
Best Local Similarity 50.0%; Pred. No. 3e-14;
Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;

QY 6 CGENEKYDCGS-KECDKKYGVDEEDDEPNVPLVRVCHQD--CVCBEGFYRNKDD 62
Db 4 CGSNERYSDCGNDKQCKERKCNEDDYKGD-----ACRSHVCERPACVCEGDFYRNKKG 58

QY 63 KCVSAEDCELDNMDFI 78
Db 59 SCVESDDCEYDNMDFI 74

RESULT 5

US-09-498-272-62
/ Sequence 62, Application US/09498272
/ Publication No. US20030113890A1
/ GENERAL INFORMATION:
/ APPLICANT: Vlasuk, George Phillip
/ Stanssens, Patrick Eric Hugo
/ Messens, Joris Hilda Lieven
/ Lauwereys, Marc Josef
/ Laroche, Yves Rene

```
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
;
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
;
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
;
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
;
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
; SEQUENCE DESCRIPTION: SEQ ID NO: 62:
US-09-498-272-62
Query Match 43.6%; Score 212; DB 11; Length 171;
Best Local Similarity 50.0%; Pred. No. 6.4e-14;
Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;
Qy 6 CGENEKDYSCGSKCEDKKCKYDGVVEEDDEPNVPCLVRYCHQD--CVCEEGFYRNKDD 62
Db 93 CGSNERYSDDCGNDKQCRKCNEDDYEKGE-----ACKSHVCERPGACVCEGDFYRNKKG 147
Qy 63 KCVSAEDCELDNMDPI 78
Db 148 SCVESDDCEYDNMDPI 163
RESULT 6
US-09-498-272-46
; Sequence 46, Application US/09498272
;
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlausk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Llieven
; Lauwereys, Marc Jozef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
;
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
;
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
;
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
;
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-498-272-46
Query Match 41.6%; Score 202; DB 11; Length 86;
Best Local Similarity 47.4%; Pred. No. 3.2e-13;
Matches 37; Conservative 15; Mismatches 22; Indels 4; Gaps 2;
Qy 5 QCGENEKDYSCGSKCEDKKCKYDGVVEEDD--PEPNVPCLVRYCHQ--DCVCEEGFYRNK 60
Db 5 KCGPGERLDCANKKPCPKCIETSEEDDDVEETDVRCLRVRCRPLKCIKDGYYRNK 64
Qy 61 DDKCVSAEDCELDNMDPI 78
```


Query Match 41.4%; Score 201; DB 11; Length 89;
Best Local Similarity 48.1%; Pred. No. 4.2e-13;

```
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-498-272-24

Query Match      41.4%; Score 201; DB 11; Length 108;
Best Local Similarity 48.18; Pred. No. 5.1e-13;
Matches 38; Conservative 13; Mismatches 24; Indels 4; Gaps 3;

Qy 3 TMOQGENEKYDSCGS-KECDKKCKYDGVBEEDDEPNVPCLVRVCHQD--CVCEEGFYRN 59
Db 26 TNACGLNEYFAECGNMKECHRCN-EEENERDEERITACLRVCFRPGACVCKDGYRN 84

Qy 60 KDKCVSAEDCELDNMF1 78
Db 85 RTGSCVBEEDDCEYENMEFI 103

RESULT 9
US-09-498-272-45
; Sequence 45, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlaauk, George Phillip
; Stanseens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440

; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-498-272-45

Query Match      40.9%; Score 199; DB 11; Length 86;
Best Local Similarity 46.2%; Pred. No. 6.4e-13;
Matches 36; Conservative 16; Mismatches 22; Indels 4; Gaps 2;

Qy 5 QCGENEKYDSCGSKECDKKCKYDGVBEEDD--EENVPCLVRVCHQ--DCVCEEGFYRN 60
Db 5 KCGPGERLDCANKKPCCKIETSEEDDDVEDTDVRLVRCBERPLKCKICKDGYRNK 64

Qy 61 DDKCVSAEDCELDNMF1 78
Db 65 KGEVTDVDCQEDFMEFI 82

RESULT 10
US-09-498-272-43
; Sequence 43, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlaauk, George Phillip
; Stanseens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
```

ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 88 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-09-498-272-43

Query Match 40.6%; Score 197.5; DB 11; Length 88;
Best Local Similarity 46.2%; Pred. No. 9.3e-13;
Matches 37; Conservative 19; Mismatches 17; Indels 7; Gaps 4;

QY 5 QCGENKYDCSGKE-CDKKCKYDGVBEEDDEPNVP---CLVRVCHQ--DCVCEGFGYR 58
Db 6 ECGPHERLD-CGNKPCERKCKIETSEEDDYEGTFRFRCLLRVCDQPVETICDDGYR 64

QY 59 NKDDKCVSAEDCELDNMDFI 78
Db 65 NKGECVTDDVCOEDFMEFI 84

RESULT 11
US-09-498-272-50
Sequence 50, Application US/09498272
Publication No. US20030113890A1
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Gansmans, Yannick Georges Jozef
Moyle, Matthew
Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,272
FILING DATE: 04-Feb-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397

FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-09-498-272-50

Query Match 40.2%; Score 195.5; DB 11; Length 84;
Best Local Similarity 50.0%; Pred. No. 1.4e-12;
Matches 40; Conservative 11; Mismatches 26; Indels 3; Gaps 3;

QY 1 KATMOGENEKYDCSGS-KECDKKCKYDGVBEEDDEPNVPCLVRVC-HQDCVCEGFGYR 58
Db 1 KSAAKCGLNKLD-CGNLKACEKCKSLDNEEDYKEDESKRSECRRCVVCDEGFGYR 59

QY 59 NKDDKCVSAEDCELDNMDFI 78
Db 60 NKGQCVTRDDCEYDNMEII 79

RESULT 12
US-09-498-272-63
Sequence 63, Application US/09498272
Publication No. US20030113890A1
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Gansmans, Yannick Georges Jozef
Moyle, Matthew
Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,272

```

1 FILING DATE: 04-Feb-2000
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: PCT/US95/13231
4 FILING DATE: October 17, 1995
5 APPLICATION NUMBER: 08/486,399
6 FILING DATE: June 5, 1995
7 APPLICATION NUMBER: 08/486,397
8 FILING DATE: June 5, 1995
9 APPLICATION NUMBER: 08/465,380
10 FILING DATE: June 5, 1995
11 APPLICATION NUMBER: 08/461,965
12 FILING DATE: June 5, 1995
13 APPLICATION NUMBER: 08/326,110
14 FILING DATE: October 18, 1994
15 ATTORNEY/AGENT INFORMATION:
16 NAME: BIGGS, SUZANNE L.
17 REGISTRATION NUMBER: 30,158
18 REFERENCE/DOCKET NUMBER: 216/270
19 TELECOMMUNICATION INFORMATION:
20 TELEPHONE: (213) 489-1600
21 TELEFAX: (213) 955-0440
22 TELEX: 67-3510
23 INFORMATION FOR SEQ ID NO: 63:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 162 amino acids
26 TYPE: amino acid
27 TOPOLOGY: linear
28 MOLECULE TYPE: peptide
29 ORIGINAL SOURCE:
30 ORGANISM: Ancylostoma caninum
31 SEQUENCE DESCRIPTION: SEQ ID NO: 63:
32 US-09-498-272-63
33
34 Query Match 40.2%; Score 195.5; DB 11; Length 162;
35 Best Local Similarity 50.0%; Pred. No. 2.8e-12;
36 Matches 40; Conservative 11; Mismatches 26; Indels 3; Gaps 3;
37
38 QY 1 KATMQCCENKDYSCGS-KECDDKKCKYDGVVEEDDEPNPVCILRVVC-HODCVCEGFYR 58
39 DB 1 KSAKKCGLNKLD-CGNLKACEKKCSLDNEEDYKEDESKRSRCSRRVCVCEGFYR 59
40
41 QY 59 NKDDKCVSAEDCELDNMDFI 78
42 DB 60 NKKGQCVRDDCEYDNMEII 79
43
44 RESULT 13
45 US-09-498-272-51
46 Sequence 51, Application US/09498272
47 Publication No. US20030113890A1
48 GENERAL INFORMATION:
49 APPLICANT: Vlaeuk, George Phillip
50 Stanssens, Patrick Eric Hugo
51 Messens, Joris Hilda Lieven
52 Lauwereys, Marc Josef
53 Laroche, Yves Rene
54 Jespers, Laurent Stephane
55 Ganssemans, Yannick Georges Jozef
56 Moyle, Matthew
57 Bergum, Peter W.
58 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
59 INHIBITORS AND ANTICOAGULANT
60 PROTEIN
61 NUMBER OF SEQUENCES: 356
62 CORRESPONDENCE ADDRESS:
63 ADDRESSEE: Lyon & Lyon
64 STREET: 633 West Fifth Street
65 Suite 4700
66 CITY: Los Angeles
67 STATE: California
68 COUNTRY: U.S.A.
69 ZIP: 90071
70 COMPUTER READABLE FORM:

```

STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
ZIP: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
FILING DATE: US/09/498,272
FILING DATE: 04-Feb-2000
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-09-498-272-64

Query Match 39.8%; Score 193.5; DB 11; Length 162;
Best Local Similarity 50.0%; Pred. No. 4.4e-12;
Matches 40; Conservative 11; Mismatches 26; Indels 3; Gaps 3;
QY 1 KATMOCGENEKYDSCGS-KECDKKCKYDGVVEEDDEPNVPCLVRC-HQDCVCEGFYR 58
Db 1 KSAKCKGLNEKLD-CGNLKACEKKCSLDNEEDYGEEDSKCRSECRICRVCCVCEGFYR 59
QY 59 NKDDKCVSAEDCELDNDFI 78
Db 60 NKKGQCVTRDDCEYDNMEII 79

RESULT 15
US-09-498-272-52
Sequence 52, Application US/09498272
Publication No. US20030113890A1
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Gansemans, Yannick Georges Jozef
Moyle, Matthew

Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
ZIP: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
FILING DATE: US/09/498,272
FILING DATE: 04-Feb-2000
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 83 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma duodenale
SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-498-272-52

Query Match 38.3%; Score 186; DB 11; Length 83;
Best Local Similarity 50.0%; Pred. No. 1.3e-11;
Matches 40; Conservative 11; Mismatches 25; Indels 4; Gaps 4;
QY 1 KATMOCGENEKYDSCGS-KECDKKCKYDGVVEEDDEPNVPCLVRC-HQDCVCEGFYR 58
Db 1 KAAKCKGLNERLD-CGNLKQCEPKCS-DLESEVEEDESCKRSECSRRVCCVCEGFYR 58
QY 59 NKDDKCVSAEDCELDNDFI 78
Db 59 NKKGKCVAKDVCEDDNMEII 78

Search completed: December 8, 2003, 09:34:03
Job time : 75.6667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 114.545 seconds
(without alignments)
116.399 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1 KATMOGENEYDCSGKEC.....VSAEDCELDNMDFIYFGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	84	20	AA1980.DAT.*
2	486	100.0	84	21	AA1981.DAT.*
3	486	100.0	91	17	AA1982.DAT.*
4	486	100.0	91	20	AA1983.DAT.*
5	486	100.0	91	20	AA1984.DAT.*
6	486	100.0	91	21	AA1985.DAT.*
7	212	43.6	82	20	AA1986.DAT.*
8	212	43.6	82	20	AA1987.DAT.*
9	212	43.6	82	21	AA1988.DAT.*

10	212	43.6	82	21	AA1989.DAT.*
11	212	43.6	171	20	AA1990.DAT.*
12	212	43.6	190	17	AA1991.DAT.*
13	212	43.6	190	20	AA1992.DAT.*
14	202	41.6	86	20	AA1993.DAT.*
15	202	41.6	86	21	AA1994.DAT.*
16	202	41.6	105	17	AA1995.DAT.*
17	202	41.6	105	20	AA1996.DAT.*
18	201	41.4	89	20	AA1997.DAT.*
19	201	41.4	108	20	AA1998.DAT.*
20	201	41.4	108	21	AA1999.DAT.*
21	199	40.9	86	20	AA2000.DAT.*
22	199	40.9	86	21	AA2001.DAT.*
23	199	40.9	105	17	AA2002.DAT.*
24	199	40.9	105	20	AA2003.DAT.*
25	197.5	40.6	88	20	AA2004.DAT.*
26	197.5	40.6	88	21	AA2005.DAT.*
27	197.5	40.6	107	17	AA2006.DAT.*
28	197.5	40.6	107	20	AA2007.DAT.*
29	195.5	40.2	84	20	AA2008.DAT.*
30	195.5	40.2	162	20	AA2009.DAT.*
31	195.5	40.2	162	21	AA2010.DAT.*
32	195.5	40.2	181	17	AA2011.DAT.*
33	195.5	40.2	181	20	AA2012.DAT.*
34	193.5	39.8	84	20	AA2013.DAT.*
35	193.5	39.8	162	20	AA2014.DAT.*
36	193.5	39.8	162	21	AA2015.DAT.*
37	193.5	39.8	181	17	AA2016.DAT.*
38	193.5	39.8	181	20	AA2017.DAT.*
39	186	38.3	83	20	AA2018.DAT.*
40	186	38.3	102	20	AA2019.DAT.*
41	186	38.3	102	21	AA2020.DAT.*
42	186	38.3	161	20	AA2021.DAT.*
43	186	38.3	161	21	AA2022.DAT.*
44	186	38.3	180	17	AA2023.DAT.*
45	184.5	38.0	87	20	AA2024.DAT.*

ALIGNMENTS

RESULT 1

AA1980.DAT.*

ID AA1980.DAT.* standard; Protein; 84 AA.

AC AA1980.DAT.*

XX 15-NOV-1999 (first entry)

DT Mature nematode extracted anticoagulant protein AcanAPc2.

DE Nematode extracted anticoagulant protein; NAP; anticoagulant;

KW serine protease inhibitor; NAP domain; factor VIIa/TF.

KW Ancylostoma caninum.

OS US5955294-A.

PN 21-SEP-1999.

PD 19-APR-1996; 96US-0634641.

PF 19-APR-1996; 96US-0634641.

PR 18-OCT-1994; 94US-0326110.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.

PR 05-JUN-1995; 95US-0486399.

PR 17-OCT-1995; 95WO-US13231.

XX (CORV-) CORVAS INT INC.

PA Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;

PI

```
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH,
PI Vlasuk GP;
XX WPI; 1999-539569/45.
XX Screening an isolated protein for Nematode-extracted Anticoagulant
XX Protein domains
XX Disclosure; Columns 142-144; 197pp; English.
XX The present sequence represents a nematode extracted anticoagulant
XX protein (NAP). The protein has activity as an anticoagulant and/or serine
XX protease inhibitor. The protein contains at least one NAP domain which
XX has selective inhibitory activity for factor VIIa/TF. The specification
XX describes a method for screening an isolated protein at least one domain
XX for factor VIIa/TF selective inhibitory activity. The method comprises
XX determining the time to clotting effected by a concentration of the
XX isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
XX activated partial thromboplastin time (aPTT) assay; calculating
XX prolongation of clotting effected by the isolated protein in each of
XX the PT and aPTT assay, with respect to a baseline clotting value for
XX each assay, where prolongation of clotting is calculated as fold
XX elevation of clotting time relative to a baseline clotting value, where
XX a doubling of clotting time is deemed a two-fold elevation; and
XX calculating a PT to aPTT prolongation ratio, where a ratio at least
XX one is indicative of factor VIIa/TF inhibitory activity. The method is
XX useful for determining if a protein has factor VIIa/TF inhibitory
XX activity.
XX SQ Sequence 84 AA;
Query Match 100.0%; Score 486; DB 20; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.7e-38;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KATWQCGNEKYDSCGSKCKKCYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
DB 1 KATWQCGNEKYDSCGSKCKKCYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
DB 61 DDKCVSAEDCELDNMDFIYPGTRN 84
RESULT 2
AAB15317
ID AAB15317 standard; Protein; 84 AA.
AC AAB15317;
XX 19-DEC-2000 (first entry)
DE A. caninum nematode-extracted anticoagulant protein AcanAPc3(mature) .
KW Nematode-extracted anticoagulant protein; AcanAPc2; blood clotting;
KW canine hookworm; thrombosis; vaccine.
XX Ancylostoma caninum.
OS US6087487-A.
PN 11-JUL-2000.
XX 12-FEB-1999; 99US-0249451.
XX 17-OCT-1995; 95WO-US13231.
PR 17-APR-1997; 97US-0809455.
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
XX PA
(CORV-) CORVAS INT INC.
PA Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;
PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;
XX WPI; 2000-531359/48.
XX N-PSDB; AAA73373.
XX New cDNA molecule encoding a protein having factor Xa inhibitory
XX activity for preventing and treating blood clotting disorders,
XX comprises nematode-extracted anticoagulant protein domains -
XX Example A; Fig 16; 197pp; English.
XX The present sequence is the Ancylostoma caninum nematode-extracted
XX anticoagulant protein AcanAPc2. Proteins of this kind have been shown to
XX be effective at preventing blood clotting without causing excessive
XX bleeding. The protein can be used in blood collection tubes to aid the
XX isolation of plasma from the blood, to prevent thrombosis which may be
XX linked to the rupture of an atherosclerotic plaque, acute myocardial
XX infarction, angina, thrombolytic therapy, percutaneous transluminal
XX coronary angioplasty, disseminated intravascular coagulopathy, infection,
XX cancer and septic shock, and to produce antibodies. In the latter
XX instance, the antibodies can be raised in order to detect infection by
XX nematodes (the coding sequence can also be used for this) or as
XX diagnostic tests. The protein can also be used as a vaccine against
XX nematode parasites.
XX SQ Sequence 84 AA;
Query Match 100.0%; Score 486; DB 21; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.7e-38;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KATWQCGNEKYDSCGSKCKKCYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
DB 1 KATWQCGNEKYDSCGSKCKKCYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
DB 61 DDKCVSAEDCELDNMDFIYPGTRN 84
RESULT 3
AAR91701
ID AAR91701 standard; Protein; 91 AA.
AC AAR91701;
XX 25-MAR-2003 (updated)
DT 17-NOV-1996 (first entry)
XX AcanAPc2.
XX AcanAP; HpoNAP; NamNAP; AceNAP; AduNAP; anticoagulant;
KW nematode-extracted anticoagulant protein; serine protease;
KW nematode; thrombosis; parasitic worm.
XX Ancylostoma caninum.
OS WO9612021-A2.
PN 25-APR-1996.
XX 17-OCT-1995; 95WO-US13231.
XX 05-JUN-1995; 95US-0486399.
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
XX PA
(CORV-) CORVAS INT INC.
```


XX Vlasuk GP, Stanssens PEH, Messens JHL, Lauwereys MJ, Laroche YR;
PI Jespers LS, Gansemans YGJ, Moyle M, Bergum PW;
XX WPI; 1996-222007/22.
DR N-PSDB; AAT12947.
XX
XX Proteins with anticoagulant and/or serine protease inhibitory
PT activity - isolated from nematodes and useful to inhibit blood
PT coagulation
XX
XX Claim 89 + 96; Fig 9; 243pp; English.
PS
XX
XX Proteins with anticoagulant and/or serine protease inhibitory
CC activity, isolated from nematodes, are useful to inhibit blood
CC coagulation. The proteins can be added to blood collection tubes
CC defining the collection of mammalian plasma. They are also useful
CC to prevent or inhibit thrombosis, and may be given alone or in
CC combination with other therapeutic or in vivo diagnostic agents.
CC The proteins can serve as immunogens to raise antibodies for use in
CC the diagnosis and identification of NAP concn. levels in biological
CC fluids, e.g. to detect mammalian infection with a parasitic worm.
CC They can also be used as immunogens in prophylactic and therapeutic
CC vaccines against parasitic worm infection. The proteins may
CC double the clotting time of human plasma in prothrombin time assays
CC when present at 10-50 nMol, and double the clotting time of human
CC plasma in activated partial thrombin time assays when present
CC at 10-100 nMol.
CC The anticoagulant proteins are pref. derived from
CC Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator
CC americanus or Heligmosomoides polygyrus.
CC The proteins pref. have 2 NAP domains and specifically inhibit
CC the catalytic activity of the factor VIIa/TF complex in the
CC presence of factor Xa or a catalytically inactive factor Xa deriv.,
CC do not specifically inhibit the activation of factor VIIa in the
CC absence of TF and do not specifically inhibit prothrombinase.
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Sequence 91 AA;
SQ
Query Match 100.0%; Score 486; DB 17; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.9e-38;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KATMQGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFYRNK 60
Db 8 KATMQGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFYRNK 67
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 68 DDKCVSAEDCELDNMDFIYPGTRN 91
RESULT 4
AAY30393
ID AAY30393 standard; Protein; 91 AA.
XX
XX AAY30393;
XX
XX 15-NOV-1999 (first entry)
XX
XX Nematode extracted anticoagulant protein AcanAPc2.
XX
XX Nematode extracted anticoagulant protein; NAP; anticoagulant;
KW serine protease inhibitor; NAP domain; factor VIIa/TF.
XX
XX Ancylostoma caninum.
OS
XX US5955294-A.
XX
XX 21-SEP-1999.
XX
XX 19-APR-1996; 96US-0634641.
XX

XX 19-APR-1996; 96US-0634641.
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
PR 17-OCT-1995; 95WO-US13231.
XX (CORV-) CORVAS INT INC.
PA
XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
PI Vlasuk GP;
XX WPI; 1999-539569/45.
DR N-PSDB; AA299999.
XX
XX Screening an isolated protein for Nematode-extracted Anticoagulant
PT Protein domains
XX
XX Example 13; Fig 9; 197pp; English.
XX
XX The present sequence represents a nematode extracted anticoagulant
CC protein (NAP). The protein has activity as an anticoagulant and/or serine
CC protease inhibitor. The protein contains at least one NAP domain which
CC has selective inhibitory activity for factor VIIa/TF. The specification
CC describes a method for screening an isolated protein at least one domain
CC for factor VIIa/TF selective inhibitory activity. The method comprises
CC isolating protein in an ex vivo prothrombin time (PT) assay and an ex vivo
CC activated partial thromboplastin time (aPTT) assay; calculating
CC prolongation of clotting effected by the isolated protein in each of
CC the PT and aPTT assay, with respect to a baseline clotting value for
CC each assay, where prolongation of clotting is calculated as fold
CC elevation of clotting time relative to a baseline clotting value, where
CC a doubling of clotting time is deemed a two-fold elevation; and
CC calculating a PT to aPTT prolongation ratio, where a ratio at least
CC one is indicative of factor VIIa/TF inhibitory activity. The method is
CC useful for determining if a protein has factor VIIa/TF inhibitory
CC activity.
XX
XX Sequence 91 AA;
SQ
Query Match 100.0%; Score 486; DB 20; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.9e-38;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KATMQGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFYRNK 60
Db 8 KATMQGENEKYDSCGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFYRNK 67
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 68 DDKCVSAEDCELDNMDFIYPGTRN 91
RESULT 5
AAY30454
ID AAY30454 standard; Protein; 91 AA.
XX
XX AAY30454;
XX
XX 15-NOV-1999 (first entry)
XX
XX Nematode extracted anticoagulant protein AcanAPc2.
XX
XX Nematode extracted anticoagulant protein; NAP; anticoagulant;
KW serine protease inhibitor; NAP domain; factor VIIa/TF.
XX
XX Ancylostoma caninum.
OS
XX US5955294-A.
XX

XX PD 21-SEP-1999.
 XX PF 19-APR-1996; 96US-0634641.
 XX PR 19-APR-1996; 96US-0634641.
 XX PR 18-OCT-1994; 94US-0326110.
 XX PR 05-JUN-1995; 95US-0461965.
 XX PR 05-JUN-1995; 95US-0465380.
 XX PR 05-JUN-1995; 95US-0486397.
 XX PR 05-JUN-1995; 95US-0486399.
 XX PR 17-OCT-1995; 95WO-US13231.
 XX PA (CORV-) CORVAS INT INC.
 XX PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
 PI Vlasuk GP;
 XX WPI; 1999-539569/45.
 XX Screening an isolated protein for Nematode-extracted Anticoagulant
 PT Protein domains
 XX Disclosure; Columns 175-176; 197pp; English.
 XX The present sequence represents a nematode extracted anticoagulant
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine
 CC protease inhibitor. The protein contains at least one NAP domain which
 CC has selective inhibitory activity for factor VIIa/TF. The specification
 CC describes a method for screening an isolated protein at least one domain
 CC for factor VIIa/TF selective inhibitory activity. The method comprises
 CC determining the time to clotting effected by a concentration of the
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
 CC activated partial thromboplastin time (APTT) assay; calculating
 CC prolongation of clotting effected by the isolated protein in each of
 CC the PT and APTT assay, with respect to a baseline clotting value for
 CC each assay, where prolongation of clotting is calculated as fold
 CC elevation of clotting time relative to a baseline clotting value, where
 CC a doubling of clotting time is deemed a two-fold elevation; and
 CC calculating a PT to APTT prolongation ratio, where a ratio at least
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is
 CC useful for determining if a protein has factor VIIa/TF inhibitory
 CC activity.
 XX SQ Sequence 91 AA;
 Query Match 100.0%; Score 486; DB 20; Length 91;
 Best Local Similarity 100.0%; Pred. No. 1.9e-38;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KATMQCGENEKYDSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
 DB 8 KATMQCGENEKYDSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 67
 QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
 DB 68 DDKCVSAEDCELDNMDFIYPGTRN 91
 RESULT 6
 AAB15346
 ID AAB15346 standard; Protein; 91 AA.
 XX AC AAB15346;
 XX DT 19-DEC-2000 (first entry)
 XX A. caninum nematode-extracted anticoagulant protein AcanAPc2.
 XX Nematode-extracted anticoagulant protein; AcanAPc2; blood clotting;
 KW canine hookworm; thrombosis; vaccine.
 XX

OS Ancylostoma caninum.
 XX US6087487-A.
 XX 11-JUL-2000.
 XX 12-FEB-1999; 99US-0249451.
 XX 17-OCT-1995; 95WO-US13231.
 XX 17-APR-1997; 97US-0809455.
 XX 18-OCT-1994; 94US-0326110.
 XX 05-JUN-1995; 95US-0461965.
 XX 05-JUN-1995; 95US-0465380.
 XX 05-JUN-1995; 95US-0486397.
 XX 05-JUN-1995; 95US-0486399.
 XX (CORV-) CORVAS INT INC.
 XX Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;
 PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;
 XX WPI; 2000-531359/48.
 XX N-PSDB; AAA73373.
 XX New cDNA molecule encoding a protein having factor Xa inhibitory
 PT activity for preventing and treating blood clotting disorders,
 PT comprises nematode-extracted anticoagulant protein domains -
 XX Disclosure; Fig 11; 197pp; English.
 XX The present sequence is the Ancylostoma caninum nematode-extracted
 CC anticoagulant protein AcanAPc2. Proteins of this kind have been shown to
 CC be effective at preventing blood clotting without causing excessive
 CC bleeding. The protein can be used in blood collection tubes to aid the
 CC isolation of plasma from the blood, to prevent thrombosis which may be
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,
 CC cancer and septic shock, and to produce antibodies. In the latter
 CC instance, the antibodies can be raised in order to detect infection by
 CC nematodes (the coding sequence can also be used for this) or as
 CC diagnostic tests. The protein can also be used as a vaccine against
 CC nematode parasites.
 XX SQ Sequence 91 AA;
 Query Match 100.0%; Score 486; DB 21; Length 91;
 Best Local Similarity 100.0%; Pred. No. 1.9e-38;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KATMQCGENEKYDSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 60
 DB 8 KATMQCGENEKYDSCGSKCEDKCKYDGVVEEDDEEPNVPCLVRVCHQDCVCEGFYRNK 67
 QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
 DB 68 DDKCVSAEDCELDNMDFIYPGTRN 91
 RESULT 7
 AAY30399
 ID AAY30399 standard; Protein; 82 AA.
 XX AC AAY30399;
 XX 15-NOV-1999 (first entry)
 XX Nematode extracted anticoagulant protein AcanAP4d2.
 XX Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX Ancylostoma ceylanicum.
 OS

DE A. ceylanicum nematode-extracted anticoagulant protein AceNAP4d2 #1.
 XX Nematode-extracted anticoagulant protein; AceNAP4d2; blood clotting;
 KW canine hookworm; thrombosis; vaccine.
 XX

OS Ancylostoma ceylanicum.

XX US6087487-A.

XX 11-JUL-2000.

XX 12-FEB-1999; 99US-0249451.

XX 17-OCT-1995; 95WO-US13231.

PR 17-APR-1997; 97US-0809455.

PR 18-OCT-1994; 94US-0326110.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.

PR 05-JUN-1995; 95US-0486399.

XX (CORV-) CORVAS INT INC.

XX Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;

PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;

XX WPI; 2000-531359/48.

XX New cDNA molecule encoding a protein having factor Xa inhibitory

PT activity for preventing and treating blood clotting disorders,

PT comprises nematode-extracted anticoagulant protein domains -

XX Disclosure; Fig 11; 197pp; English.

XX The present sequence is the Ancylostoma ceylanicum nematode-extracted
 CC anticoagulant protein AceNAP4d2. Proteins of this kind have been shown to
 CC be effective at preventing blood clotting without causing excessive
 CC bleeding. The protein can be used in blood collection tubes to aid the
 CC isolation of plasma from the blood, to prevent thrombosis which may be
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,
 CC cancer and septic shock, and to produce antibodies. In the latter
 CC instance, the antibodies can be raised in order to detect infection by
 CC nematodes or as diagnostic tests. The protein can also be used as a
 CC vaccine against nematode parasites.

XX Sequence 82 AA;

Query Match 43.6%; Score 212; DB 21; Length 82;

Best Local Similarity 50.0%; Pred. No. 9.1e-13;

Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;

QY 6 CGENKDYSCGS-KECDKKCKYDGVVEEDDEEPNVPCILVRVCHQD--CVCEEGFYRNKDD 62

DB 4 CGSNERYSDCGNDKQCKERNEDDYKGDGDE-----ACRSHVCERPACVCEGDFYRNKKG 58

QY 63 KCVSAEDCELDNMDFI 78

DB 59 SCVESDDCEYDNMDFI 74

RESULT 10

AAB15307

ID AAB15307 standard; Protein; 82 AA.

XX AAB15307;

XX 19-DEC-2000 (first entry)

DE A. ceylanicum nematode-extracted anticoagulant protein AceNAP4d2 #2.

XX Nematode-extracted anticoagulant protein; AceNAP4d2; blood clotting;

KW canine hookworm; thrombosis; vaccine.

XX Ancylostoma ceylanicum.

XX US6087487-A.

XX 11-JUL-2000.

XX 12-FEB-1999; 99US-0249451.

XX 17-OCT-1995; 95WO-US13231.

PR 17-APR-1997; 97US-0809455.

PR 18-OCT-1994; 94US-0326110.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.

PR 05-JUN-1995; 95US-0486399.

XX (CORV-) CORVAS INT INC.

XX Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;

PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;

XX WPI; 2000-531359/48.

XX New cDNA molecule encoding a protein having factor Xa inhibitory

PT activity for preventing and treating blood clotting disorders,

PT comprises nematode-extracted anticoagulant protein domains -

XX Disclosure; Fig 16; 197pp; English.

XX The present sequence is the Ancylostoma ceylanicum nematode-extracted
 CC anticoagulant protein AceNAP4d2. Proteins of this kind have been shown to
 CC be effective at preventing blood clotting without causing excessive
 CC bleeding. The protein can be used in blood collection tubes to aid the
 CC isolation of plasma from the blood, to prevent thrombosis which may be
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,
 CC cancer and septic shock, and to produce antibodies. In the latter
 CC instance, the antibodies can be raised in order to detect infection by
 CC nematodes or as diagnostic tests. The protein can also be used as a
 CC vaccine against nematode parasites.

XX Sequence 82 AA;

Query Match 43.6%; Score 212; DB 21; Length 82;

Best Local Similarity 50.0%; Pred. No. 9.1e-13;

Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;

QY 6 CGENKDYSCGS-KECDKKCKYDGVVEEDDEEPNVPCILVRVCHQD--CVCEEGFYRNKDD 62

DB 4 CGSNERYSDCGNDKQCKERNEDDYKGDGDE-----ACRSHVCERPACVCEGDFYRNKKG 58

QY 63 KCVSAEDCELDNMDFI 78

DB 59 SCVESDDCEYDNMDFI 74

RESULT 11

AAY30435

ID AAY30435 standard; Protein; 171 AA.

XX AAY30435;

XX 15-NOV-1999 (first entry)

DE Mature nematode extracted anticoagulant protein AceNAP4.

XX Nematode extracted anticoagulant protein; NAP; anticoagulant;

KW serine protease inhibitor; NAP domain; factor VIIa/TF.

XX Ancylostoma ceylanicum.

XX US955294-A.
 XX 21-SEP-1999.
 XX 19-APR-1996; 96US-0634641.
 XX 19-APR-1996; 96US-0634641.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.
 PR 05-JUN-1995; 95US-0486397.
 PR 05-JUN-1995; 95US-0486399.
 PR 17-OCT-1995; 95WO-US13231.
 XX (CORV-) CORVAS INT INC.
 XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
 PI Vlausk GP;
 XX WPI; 1999-539569/45.
 XX Screening an isolated protein for Nematode-extracted Anticoagulant
 PT Protein domains
 XX Disclosure; Fig 17; 197pp; English.
 XX The present sequence represents a nematode extracted anticoagulant
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine
 CC protease inhibitor. The protein contains at least one NAP domain which
 CC has selective inhibitory activity for factor VIIa/TF. The specification
 CC describes a method for screening an isolated protein at least one domain
 CC for factor VIIa/TF selective inhibitory activity. The method comprises
 CC determining the time to clotting effected by a concentration of the
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
 CC activated partial thromboplastin time (aPTT) assay; calculating
 CC the PT and aPTT assay, with respect to a baseline clotting value for
 CC each assay, where prolongation of clotting is calculated as fold
 CC elevation of clotting time relative to a baseline clotting value, where
 CC a doubling of clotting time is deemed a two-fold elevation; and
 CC calculating a PT to aPTT prolongation ratio, where a ratio at least
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is
 CC useful for determining if a protein has factor VIIa/TF inhibitory
 CC activity.
 XX Sequence 171 AA;
 SQ
 Query Match 43.6%; Score 212; DB 20; Length 171;
 Best Local Similarity 50.0%; Pred. No. 1.9e-12;
 Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;
 QY 6 CGENEKDYSCGS-KCEDKKCKYDGVVEEDDEBNVFLVRVCHQD--CVCSEGFYRNKDD 62
 DB 93 CGSNERYSDCGNDKQCKERCNEDDYKGDGDE-----ACRSHVCERPGACVCEGDFYRNKKG 147
 QY 63 KCVSAEDCELDNNDFI 78
 DB 148 SCVESDCCBYDNDNFI 163
 RESULT 12
 AAR91710
 ID AAR91710 standard; Protein; 190 AA.
 XX
 AC AAR91710;
 XX
 DT 25-MAR-2003 (updated)
 DT 17-NOV-1996 (first entry)
 XX
 XX AcenAP4.
 DE
 XX

KW AcenAP; HpoNAP; NamNAP; AceNAP; AdunAP; antioagulant;
 KW nematode-extracted anticoagulant protein; serine protease;
 KW nematode; thrombosis; parasitic worm.
 XX
 OS Ancylostoma ceylanicum.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..108
 FT /label= AceNAP4d1
 FT Domain 109..190
 FT /label= AceNAP4d2
 XX
 XX W09612021-A2.
 XX
 XX 25-APR-1996.
 XX
 XX 17-OCT-1995; 95WO-US13231.
 XX
 XX 05-JUN-1995; 95US-0486399.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.
 PR 05-JUN-1995; 95US-0486397.
 XX (CORV-) CORVAS INT INC.
 XX
 XX Vlausk GP, Stanssens PEH, Messens JHL, Lauwereys MJ, Laroche YR;
 PI Jespers LS, Gansemans YGJ, Moyle M, Bergum PW;
 XX WPI; 1996-222007/22.
 N-PSDB; AAT12956.
 XX Proteins with anticoagulant and/or serine protease inhibitory
 PT activity - isolated from nematodes and useful to inhibit blood
 PT coagulation
 XX Claim 221; Fig 7A; 243pp; English.
 CC Proteins with anticoagulant and/or serine protease inhibitory
 CC activity, isolated from nematodes, are useful to inhibit blood
 CC coagulation. The proteins can be added to blood collection tubes
 CC defining the collection of mammalian plasma. They are also useful
 CC to prevent or inhibit thrombosis, and may be given alone or in
 CC combination with other therapeutic or in vivo diagnostic agents.
 CC The proteins can serve as immunogens to raise antibodies for use in
 CC the diagnosis and identification of NAP concn. levels in biological
 CC fluids, e.g. to detect mammalian infection with a parasitic worm.
 CC They can also be used as immunogens in prophylactic and therapeutic
 CC vaccines against parasitic worm infection. The proteins may
 CC double the clotting time of human plasma in prothrombin time assays
 CC when present at 10-50 nMol, and double the clotting time of human
 CC plasma in activated partial thrombin time assays when present
 CC at 10-100 nMol.
 CC The anticoagulant proteins are pref. derived from
 CC Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator
 CC americanus or Heligmosomoides polygyrus
 CC The proteins pref. have 2 NAP domains and specifically inhibit
 CC the catalytic activity of the factor VIIa/TF complex in the
 CC presence of factor Xa or a catalytically inactive factor Xa deriv.,
 CC do not specifically inhibit the activation of factor VIIa in the
 CC absence of TF and do not specifically inhibit prothrombinase.
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 XX Sequence 190 AA;
 SQ

Query Match 43.6%; Score 212; DB 17; Length 190;
 Best Local Similarity 50.0%; Pred. No. 2.1e-12;
 Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;

QY 6 CGENEKDYSCGS-KCEDKKCKYDGVVEEDDEBNVFLVRVCHQD--CVCSEGFYRNKDD 62
 DB 112 CGSNERYSDCGNDKQCKERCNEDDYKGDGDE-----ACRSHVCERPGACVCEGDFYRNKKG 166

QY 63 KCVSAEDCELDNMDFI 78
 ||:|||||
 Db 167 SCVESDDCEYDNMDFI 182

RESULT 13
 AAY30384
 ID AAY30384 standard; Protein; 190 AA.

AC AAY30384;
 XX
 DT 15-NOV-1999 (first entry)
 XX
 DE Nematode extracted anticoagulant protein AcenAP4.
 XX
 KW Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX
 OS Ancylostoma ceylanicum.

XX US9595294-A.

XX 21-SEP-1999.

XX 19-APR-1996; 96US-0634641.

XX 19-APR-1996; 96US-0634641.

PR 18-OCT-1994; 94US-0326110.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.

PR 05-JUN-1995; 95US-0486399.

PR 17-OCT-1995; 95WO-US13231.

XX (CORV-) CORVAS INT INC.

XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;

PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;

PI Vlasuk GP;

XX WPI; 1999-539569/45.

DR NP-SDB; AAZ10441.

XX Screening an isolated protein for Nematode-extracted Anticoagulant

PT Protein domains

PS Example 9; Fig 7A; 197pp; English.

XX The present sequence represents a nematode extracted anticoagulant
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine
 CC protease inhibitor. The protein contains at least one NAP domain which
 CC has selective inhibitory activity for factor VIIa/TF. The specification
 CC describes a method for screening an isolated protein at least one domain
 CC for factor VIIa/TF selective inhibitory activity. The method comprises
 CC determining the time to clotting effected by a concentration of the
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
 CC activated partial thromboplastin time (aPTT) assay; calculating
 CC prolongation of clotting effected by the isolated protein in each of
 CC the PT and aPTT assay, with respect to a baseline clotting value for
 CC each assay, where prolongation of clotting is calculated as fold
 CC elevation of clotting time relative to a baseline clotting value, where
 CC a doubling of clotting time is deemed a two-fold elevation; and
 CC calculating a PT to aPTT prolongation ratio, where a ratio at least
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is
 CC useful for determining if a protein has factor VIIa/TF inhibitory
 CC activity.

XX Sequence 190 AA;

Query Match 43.6%; Score 212; DB 20; Length 190;
 Best Local Similarity 50.0%; Pred. No. 2.1e-12;
 Matches 38; Conservative 11; Mismatches 19; Indels 8; Gaps 3;

QY 6 CGENKDYSGS-KCKDKKYDGVVEEDDEBNVPCVLRVCHOD--CVCEGFYRNKDD 62
 ||:|||||
 Db 112 CGSNERYSDGNDKQCKERCKNEDDYKGB-----ACRSHVCPGACVCEDFYNNKG 166

QY 63 KCVSAEDCELDNMDFI 78
 ||:|||||
 Db 167 SCVESDDCEYDNMDFI 182

RESULT 14

AAY30419

ID AAY30419 standard; Protein; 86 AA.

XX AAY30419;

XX 15-NOV-1999 (first entry)

XX Mature nematode extracted anticoagulant protein AcanAP44.

XX Nematode extracted anticoagulant protein; NAP; anticoagulant;

XX serine protease inhibitor; NAP domain; factor VIIa/TF.

XX Ancylostoma caninum.

XX US9595294-A.

XX 21-SEP-1999.

XX 19-APR-1996; 96US-0634641.

XX 19-APR-1996; 96US-0634641.

PR 18-OCT-1994; 94US-0326110.

PR 05-JUN-1995; 95US-0461965.

PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.

PR 05-JUN-1995; 95US-0486399.

PR 17-OCT-1995; 95WO-US13231.

XX (CORV-) CORVAS INT INC.

XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;

PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;

PI Vlasuk GP;

XX WPI; 1999-539569/45.

DR Screening an isolated protein for Nematode-extracted Anticoagulant

XX Protein domains

PS Disclosure; Columns 131-134; 197pp; English.

XX The present sequence represents a nematode extracted anticoagulant

CC protein (NAP). The protein has activity as an anticoagulant and/or serine

CC protease inhibitor. The protein contains at least one NAP domain which

CC describes a method for screening an isolated protein at least one domain

CC for factor VIIa/TF selective inhibitory activity. The method comprises

CC determining the time to clotting effected by a concentration of the

CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo

CC activated partial thromboplastin time (aPTT) assay; calculating

CC prolongation of clotting effected by the isolated protein in each of

CC the PT and aPTT assay, with respect to a baseline clotting value for

CC each assay, where prolongation of clotting is calculated as fold

CC elevation of clotting time relative to a baseline clotting value, where

CC a doubling of clotting time is deemed a two-fold elevation; and

CC calculating a PT to aPTT prolongation ratio, where a ratio at least

CC one is indicative of factor VIIa/TF inhibitory activity. The method is

CC useful for determining if a protein has factor VIIa/TF inhibitory

CC activity.

XX Sequence 86 AA;

Query Match 41.6%; Score 202; DB 20; Length 86;

Best Local Similarity 47.4%; Pred. No. 8.3e-12;
Matches 37; Conservative 15; Mismatches 22; Indels 4; Gaps 2;

QY 5 QCGENKYSKSGSKCKKCKYGVREEDD--EENVPCLVRVCHQ--DCVCEEGFYRNK 60
Db 5 KCGPGERLDCANKKPCPKIETSEEDDDVEETDVRCLVRVCRPLKCKICKDGYRNK 64
QY 61 DDKCVSAEDCELDNMDPI 78
Db 65 KGECVTDDVCQEDFMEFI 82

RESULT 15

AAB15304
ID AAB15304 standard; Protein; 86 AA.

XX AAB15304;

DT 19-DEC-2000 (first entry)

XX A. caninum nematode-extracted anticoagulant protein AcanAP44.

KW Nematode-extracted anticoagulant protein; AcanAP44; blood clotting;
KW canine hookworm; thrombosis; vaccine.

XX Ancylostoma caninum.

XX US6087487-A.

XX 11-JUL-2000.

XX 12-FEB-1999; 99US-0249451.

XX 17-OCT-1995; 95WO-US13231.

XX 17-APR-1997; 97US-0809455.

XX 18-OCT-1994; 94US-0326110.

XX 05-JUN-1995; 95US-0461965.

XX 05-JUN-1995; 95US-0465380.

XX 05-JUN-1995; 95US-0486397.

XX 05-JUN-1995; 95US-0486399.

DR WPI; 2000-531359/48.

DR N-PSDB; AAA73379.

XX New cDNA molecule encoding a protein having factor Xa inhibitory

PT activity for preventing and treating blood clotting disorders,

PT comprises nematode-extracted anticoagulant protein domains -

XX Disclosure; Fig 16; 197pp; English.

XX The present sequence is the Ancylostoma caninum nematode-extracted
XX anticoagulant protein AcanAP44. Proteins of this kind have been shown to
XX be effective at preventing blood clotting without causing excessive
XX bleeding. The protein can be used in blood collection tubes to aid the
XX isolation of plasma from the blood, to prevent thrombosis which may be
XX linked to the rupture of an atherosclerotic plaque, acute myocardial
XX infarction, angina, thrombolytic therapy, percutaneous transluminal
XX coronary angioplasty, disseminated intravascular coagulopathy, infection,
XX cancer and septic shock, and to produce antibodies. In the latter
XX instance, the antibodies can be raised in order to detect infection by
XX nematodes (the coding sequence can also be used for this) or as
XX diagnostic tests. The protein can also be used as a vaccine against
XX nematode parasites.

XX Sequence 86 AA;

Query Match 41.6%; Score 202; DB 21; Length 86;

Best Local Similarity 47.4%; Pred. No. 8.3e-12;

Matches 37; Conservative 15; Mismatches 22; Indels 4; Gaps 2;

QY 5 QCGENKYSKSGSKCKKCKYGVREEDD--EENVPCLVRVCHQ--DCVCEEGFYRNK 60
Db 5 KCGPGERLDCANKKPCPKIETSEEDDDVEETDVRCLVRVCRPLKCKICKDGYRNK 64
QY 61 DDKCVSAEDCELDNMDPI 78
Db 65 KGECVTDDVCQEDFMEFI 82

Search completed: December 8, 2003, 09:16:38
Job time : 114.545 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:31 ; Search time 39.0303 Seconds
(without alignments)
91.060 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486

Sequence: 1 KATMOCGENEKDYSCGSKC.....VSAEDCELDNMDFIYPGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	100.0	84	2	US-08-465-380-59
2	486	100.0	84	2	US-08-486-397-59
3	486	100.0	84	2	US-08-486-399-59
4	486	100.0	84	2	US-08-461-965-59
5	486	100.0	84	2	US-08-634-641-59
6	486	100.0	84	3	US-09-249-471-59
7	486	100.0	84	3	US-09-249-472-59
8	486	100.0	84	3	US-09-249-451-59
9	486	100.0	84	3	US-08-809-455-59
10	486	100.0	84	3	US-09-249-461-59
11	486	100.0	84	3	US-09-249-448-59
12	486	100.0	84	4	US-09-249-473-59
13	486	100.0	91	2	US-08-465-380-128
14	486	100.0	91	2	US-08-480-478-50
15	486	100.0	91	2	US-08-486-397-128
16	486	100.0	91	2	US-08-486-399-128
17	486	100.0	91	2	US-08-461-965-128
18	486	100.0	91	2	US-08-326-110A-50
19	486	100.0	91	2	US-08-634-641-128
20	486	100.0	91	3	US-09-249-471-128
21	486	100.0	91	3	US-09-249-472-128
22	486	100.0	91	3	US-09-249-451-128
23	486	100.0	91	3	US-08-809-455-128
24	486	100.0	91	3	US-09-249-461-128
25	486	100.0	91	3	US-09-249-448-128
26	486	100.0	91	4	US-09-249-473-128
27	212	43.6	82	2	US-08-465-380-25

28	212	43.6	82	2	US-08-465-380-49	Sequence 49, Appl
29	212	43.6	82	2	US-08-480-478-54	Sequence 54, Appl
30	212	43.6	82	2	US-08-486-397-25	Sequence 25, Appl
31	212	43.6	82	2	US-08-486-397-49	Sequence 49, Appl
32	212	43.6	82	2	US-08-486-399-25	Sequence 25, Appl
33	212	43.6	82	2	US-08-486-399-49	Sequence 49, Appl
34	212	43.6	82	2	US-08-461-965-25	Sequence 25, Appl
35	212	43.6	82	2	US-08-461-965-49	Sequence 49, Appl
36	212	43.6	82	2	US-08-326-110A-54	Sequence 54, Appl
37	212	43.6	82	2	US-08-634-641-25	Sequence 25, Appl
38	212	43.6	82	2	US-08-634-641-49	Sequence 49, Appl
39	212	43.6	82	3	US-09-249-471-25	Sequence 25, Appl
40	212	43.6	82	3	US-09-249-471-49	Sequence 49, Appl
41	212	43.6	82	3	US-09-249-472-25	Sequence 25, Appl
42	212	43.6	82	3	US-09-249-472-49	Sequence 49, Appl
43	212	43.6	82	3	US-09-249-451-25	Sequence 25, Appl
44	212	43.6	82	3	US-09-249-451-49	Sequence 49, Appl
45	212	43.6	82	3	US-08-809-455-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1

US-08-465-380-59
; Sequence 59, Application US/08465380
; Patent No 5863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NENATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,380
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum

US-08-461-965-59
; Sequence 59, Application US/08461965
; Patent No. 5872098
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,965
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 210/243
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-461-965-59

Query Match 100.0%; Score 486; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KATMOGGENEYDSCGSKCKKDYGVVEEDDEPNVPCLVRVCHQDCVCEGFRNK-60
Db 1 KATMOGGENEYDSCGSKCKKDYGVVEEDDEPNVPCLVRVCHQDCVCEGFRNK 60

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 5
US-08-634-641-59
; Sequence 59, Application US/08634641
; Patent No. 5955294
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George P. Vlasuk
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Mensens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef

; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew W.
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/634,641
; FILING DATE: April 19, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 219/136
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-634-641-59

Query Match 100.0%; Score 486; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMOGGENEYDSCGSKCKKDYGVVEEDDEPNVPCLVRVCHQDCVCEGFRNK 60
Db 1 KATMOGGENEYDSCGSKCKKDYGVVEEDDEPNVPCLVRVCHQDCVCEGFRNK 60
QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 6
US-09-249-471-59
; Sequence 59, Application US/09249471
; Patent No. 6040441

```

; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Berghum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,471
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-09-249-471-59
;
; Query Match 100.0%; Score 486; DB 3; Length 84;
; Best Local Similarity 100.0%; Pred. No. 4e-43;
; Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 KATMCQGNENKYDSCGSKCKYKYGVEBDDDEPNVPCLVRVCHODCVCEGFFYRNK 60
; DB 1 KATMCQGNENKYDSCGSKCKYKYGVEBDDDEPNVPCLVRVCHODCVCEGFFYRNK 60
;
; QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
;
; Query Match 100.0%; Score 486; DB 3; Length 84;
; Best Local Similarity 100.0%; Pred. No. 4e-43;
; Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 KATMCQGNENKYDSCGSKCKYKYGVEBDDDEPNVPCLVRVCHODCVCEGFFYRNK 60
; DB 1 KATMCQGNENKYDSCGSKCKYKYGVEBDDDEPNVPCLVRVCHODCVCEGFFYRNK 60
;
; QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
;
; RESULT 7
; US-09-249-472-59
; Sequence 59, Application US/09249472
; Patent No. 6046318
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Berghum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,472
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-09-249-472-59
;
; Query Match 100.0%; Score 486; DB 3; Length 84;
; Best Local Similarity 100.0%; Pred. No. 4e-43;
; Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 KATMCQGNENKYDSCGSKCKYKYGVEBDDDEPNVPCLVRVCHODCVCEGFFYRNK 60
; DB 1 KATMCQGNENKYDSCGSKCKYKYGVEBDDDEPNVPCLVRVCHODCVCEGFFYRNK 60
;
; QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
;
; DB 61 DDKCVSAEDCELDNMDFIYPGTRN 84
```

Best Local Similarity 100.0%; Pred. No. 4e-43; Mismatches 0; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFRNK 60
Db 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFRNK 60

QY 61 DDKVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKVSAEDCELDNMDFIYPGTRN 84

RESULT 8
US-09-249-451-59
; Sequence 59, Application US/09249451
; Patent No. 6087487
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Ganssems, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,451
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/809,455
FILING DATE: April 17, 1997
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
US-09-249-451-59

Query Match 100.0%; Score 486; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 4e-43; Mismatches 0; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFRNK 60
Db 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCCEGFRNK 60

QY 61 DDKVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKVSAEDCELDNMDFIYPGTRN 84

RESULT 9
US-08-809-455-59
; Sequence 59, Application US/08809455
; Patent No. 6090916
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Ganssems, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,455
FILING DATE: April 17, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-809-455-59

Query Match 100.0%; Score 486; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDCSGKCEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60
Db 1 KATMQCGENEKYDCSGKCEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60

Qy 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 10
US-09-249-461-59
; Sequence 59, Application US/09249461
; Patent No. 6096877
; GENERAL INFORMATION:
; APPLICANT: Vlaeuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,461
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
```

```
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-09-249-461-59

Query Match 100.0%; Score 486; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDCSGKCEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60
Db 1 KATMQCGENEKYDCSGKCEDKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCEGFYRNK 60

Qy 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 11
US-09-249-448-59
; Sequence 59, Application US/09249448
; Patent No. 6121435
; GENERAL INFORMATION:
; APPLICANT: Vlaeuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,448
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
```

```

; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
;
US-09-249-448-59

Query Match 100.0%; Score 486; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCVEEGFYRNK 60
Db 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCVEEGFYRNK 60

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 12
US-09-249-473-59
; Sequence 59, Application US/09249473
; Patent No. 6534629
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansmans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0

```

```

; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,473
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
;
US-09-249-473-59

Query Match 100.0%; Score 486; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 4e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCVEEGFYRNK 60
Db 1 KATMQGENEKYDSCGSKCKKYDGVVEEDDEPNVPCLVRVCHQDVCVEEGFYRNK 60

QY 61 DDKCVSAEDCELDNMDFIYPGTRN 84
Db 61 DDKCVSAEDCELDNMDFIYPGTRN 84

RESULT 13
US-08-465-380-128
; Sequence 128, Application US/08465380
; Patent No. 5863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansmans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

```

```
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,380
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-465-380-128

Query Match 100.0%; Score 486; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 4.3e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDCSGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCBEGFYRNK 60
Db 8 KATMQCGENEKYDCSGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCBEGFYRNK 67

Qy 61 DDKCVSAEDCCLDNMDFIYPGTRN 84
Db 68 DDKCVSAEDCCLDNMDFIYPGTRN 91

RESULT 14
US-08-480-478-50
; Sequence 50, Application US/08480478
; Patent No. 5864009
; GENERAL INFORMATION:
; APPLICANT: GEORGE P. VLASUK; PATRICK ERIC
; APPLICANT: HUGO STANSSENS; JORIS HILDA
; APPLICANT: LIEVEN MESSENS; MARC JOZEF
; APPLICANT: LAUWEREYS; YVES RENE LAROCHE;
; APPLICANT: LAURENT STEPHANE JESPEERS; and
; APPLICANT: YANNICK GEORGES JOZEF
; APPLICANT: GANSEMAN
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTI-
; TITLE OF INVENTION: CONGLUTIN PROTEIN
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,478
```

```
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: 18 OCTOBER 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 208/290
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 91 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-480-478-50

Query Match 100.0%; Score 486; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 4.3e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KATMQCGENEKYDCSGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCBEGFYRNK 60
Db 8 KATMQCGENEKYDCSGSKCKKCKYDGVVEEDDEPNVPCLVRVCHQDCVCBEGFYRNK 67

Qy 61 DDKCVSAEDCCLDNMDFIYPGTRN 84
Db 68 DDKCVSAEDCCLDNMDFIYPGTRN 91

RESULT 15
US-08-486-397-128
; Sequence 128, Application US/08486397
; Patent No. 5866542
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 357
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,397
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/269
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
```


TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 128:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-128

Query Match 100.0%; Score 486; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 4.3e-43;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KATWOCGENEKYDSCGSKCKKCKYDGVVEEDDEEPNVFCLYRVCHQDCVCEGFYRNK 60
Db 8 KATWOCGENEKYDSCGSKCKKCKYDGVVEEDDEEPNVFCLYRVCHQDCVCEGFYRNK 67
QY 61 DDKVSAEDCELDNMDFIYFGTRN 84
Db 68 DDKVSAEDCELDNMDFIYFGTRN 91

Search completed: December 8, 2003, 09:20:28
Job time : 39.0303 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:18:46 ; Search time 6.22222 Seconds
(without alignments)
209.232 Million cell updates/sec

Title: US-09-498-556C-79
Perfect score: 32
Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	5	11	US-09-498-272-70
2	30	93.8	7	11	US-09-498-272-79
3	30	93.8	78	11	US-09-498-272-47
4	30	93.8	82	11	US-09-498-272-25
5	30	93.8	82	11	US-09-498-272-49
6	30	93.8	83	11	US-09-498-272-52
7	30	93.8	84	11	US-09-498-272-50
8	30	93.8	84	11	US-09-498-272-51
9	30	93.8	84	11	US-09-498-272-59
10	30	93.8	89	11	US-09-498-272-48
11	30	93.8	91	11	US-09-498-272-128
12	30	93.8	102	11	US-09-498-272-27
13	30	93.8	108	11	US-09-498-272-24
14	30	93.8	113	15	US-10-156-761-11215
15	30	93.8	138	10	US-09-764-877-1934

16	30	93.8	161	11	US-09-498-272-65
17	30	93.8	162	11	US-09-498-272-63
18	30	93.8	162	11	US-09-498-272-64
19	30	93.8	171	11	US-09-498-272-62
20	30	93.8	260	10	US-09-738-626-3828
21	30	93.8	268	10	US-09-864-866-43
22	30	93.8	475	12	US-10-205-219-65
23	30	93.8	533	15	US-10-207-655-75
24	30	93.8	1564	10	US-09-801-368-244
25	27	84.4	5	11	US-09-498-272-72
26	27	84.4	7	11	US-09-498-272-81
27	27	84.4	79	9	US-09-864-761-45609
28	27	84.4	86	11	US-09-498-272-45
29	27	84.4	86	11	US-09-498-272-46
30	27	84.4	87	11	US-09-498-272-44
31	27	84.4	88	11	US-09-498-272-43
32	27	84.4	173	9	US-09-925-302-855
33	27	84.4	260	10	US-09-858-546-5
34	27	84.4	392	10	US-09-813-718-16
35	27	84.4	415	10	US-09-813-718-14
36	27	84.4	437	10	US-09-813-718-12
37	27	84.4	471	11	US-09-919-039-163
38	27	84.4	471	12	US-10-247-671-166
39	27	84.4	471	15	US-10-126-467B-2
40	27	84.4	475	9	US-09-925-302-558
41	27	84.4	484	10	US-09-813-718-10
42	27	84.4	530	10	US-09-858-546-2
43	27	84.4	530	14	US-10-108-605-171
44	27	84.4	2771	10	US-09-808-602-82
45	27	84.4	2771	11	US-09-800-198-70

ALIGNMENTS

RESULT 1

US-09-498-272-70
; Sequence 70, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Ganssems, Yannick Georges Jozef
Moyle, Matthew
Berghum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN

NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700

STATE: Los Angeles
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/498,272
FILING DATE: 04-Feb-2000

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995

APPLICATION NUMBER: 08/486.399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486.397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465.380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461.965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326.110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
SEQUENCE DESCRIPTION: SEQ ID NO: 70:

Query Match 93.8%; Score 30; DB 11; Length 5;
Best Local Similarity 100.0%; Pred. No. 6.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 1 GFYRN 5

RESULT 2

US-09-498-272-79
Sequence 79, Application US/09498272
Publication No. US20030113890A1
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Gansmans, Yannick Georges Jozef
Moyle, Matthew
Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER READABLE FORM:
Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,272
FILING DATE: 04-Feb-2000
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486.399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486.397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465.380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461.965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326.110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
FEATURE:
OTHER INFORMATION: Xaa in locations 1 and 2
is an amino acid, provided
that at least one Xaa is Glu or
Asp.
SEQUENCE DESCRIPTION: SEQ ID NO: 79:

US-09-498-272-79

Query Match 93.8%; Score 30; DB 11; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 3 GFYRN 7

RESULT 3

US-09-498-272-47
Sequence 47, Application US/09498272
Publication No. US20030113890A1
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Gansmans, Yannick Georges Jozef
Moyle, Matthew
Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
SUITE 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER READABLE FORM:
storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,272
FILING DATE: 04-Feb-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 78 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-498-272-47

Query Match 93.8%; Score 30; DB 11; Length 78;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 54 GFYRN 58

RESULT 4

US-09-498-272-25
Sequence 25, Application US/09498272
Publication No. US20030113890A1

GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jaspers, Laurent Stephane
Ganssemans, Yannick Georges Jozef
Moyle, Matthew
Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT

PROTEIN

NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street

Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,272
FILING DATE: 04-Feb-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma ceylanicum
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-498-272-25

Query Match 93.8%; Score 30; DB 11; Length 82;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 51 GFYRN 55

RESULT 5

US-09-498-272-49

Sequence 49, Application US/09498272
Publication No. US20030113890A1

GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jaspers, Laurent Stephane
Ganssemans, Yannick Georges Jozef
Moyle, Matthew
Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT

PROTEIN

NUMBER OF SEQUENCES: 356

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

Suite 4700

CITY: Los Angeles

STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,272
FILING DATE: 04-Feb-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma ceylanicum
SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-09-498-272-49

Query Match 93.8%; Score 30; DB 11; Length 82;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 51 GFYRN 55

RESULT 6

US-09-498-272-52
Sequence 52, Application US/09498272
Publication No. US20030113890A1
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Laureys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Gansmans, Yannick Georges Jozef
Moyle, Matthew
Berghum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street
Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,272
FILING DATE: 04-Feb-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 83 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma duodenale
SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-498-272-52

Query Match 93.8%; Score 30; DB 11; Length 83;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 55 GFYRN 59

RESULT 7

US-09-498-272-50
Sequence 50, Application US/09498272
Publication No. US20030113890A1
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Laureys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Gansmans, Yannick Georges Jozef
Moyle, Matthew
Berghum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN

NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
CITY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,272
FILING DATE: 04-Feb-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
SEQUENCE DESCRIPTION: SEQ ID NO: 50:

Query Match 93.8%; Score 30; DB 11; Length 84;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 56 GFYRN 60

RESULT 8
US-09-498-272-51
; Sequence 51, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.

TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
CITY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,272
FILING DATE: 04-Feb-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
SEQUENCE DESCRIPTION: SEQ ID NO: 51:

US-09-498-272-51

Query Match 93.8%; Score 30; DB 11; Length 84;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 56 GFYRN 60

RESULT 9
US-09-498-272-59
; Sequence 59, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane

; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; ORGANISM: Ancylostoma caninum
; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
US-09-498-272-59

Query Match 93.8%; Score 30; DB 11; Length 84;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 55 GFYRN 59
|||||

RESULT 10
US-09-498-272-48
; Sequence 48, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Scansens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven

; Lauwereys, Marc Jozef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemans, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
; SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-09-498-272-48

Query Match 93.8%; Score 30; DB 11; Length 89;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 61 GFYRN 65
|||||

RESULT 11
US-09-498-272-128
; Sequence 128, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:

APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Gansemans, Yannick Georges Jozef
Moyle, Matthew
Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN

NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,272
FILING DATE: 04-Feb-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 128:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
SEQUENCE DESCRIPTION: SEQ ID NO: 128:

US-09-498-272-128
Query Match 93.8%; Score 30; DB 11; Length 91;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
|||||

Db 62 GFYRN 66

RESULT 12

US-09-498-272-128

Sequence 27, Application US/09498272
Publication No. US20030113890A1
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
Stanssens, Patrick Eric Hugo
Messens, Joris Hilda Lieven
Lauwereys, Marc Josef
Laroche, Yves Rene
Jespers, Laurent Stephane
Gansemans, Yannick Georges Jozef
Moyle, Matthew
Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT
PROTEIN

NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/498,272
FILING DATE: 04-Feb-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma duodenale
SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-498-272-27
Query Match 93.8%; Score 30; DB 11; Length 102;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
|||||

Db 74 GFYRN 78

RESULT 13
US-09-498-272-24
; Sequence 24, Application US/09498272
; Publication No. US20030113890A1
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; Stanssens, Patrick Eric Hugo
; Messens, Joris Hilda Lieven
; Lauwereys, Marc Josef
; Laroche, Yves Rene
; Jespers, Laurent Stephane
; Gansemaus, Yannick Georges Jozef
; Moyle, Matthew
; Bergum, Peter W.
; TITLE OF INVENTION: NENATODE-EXTRACTED SERINE PROTEASE
; INHIBITORS AND ANTICOAGULANT
; PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/498,272
; FILING DATE: 04-Feb-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/POCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma ceylanicum
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-498-272-24

Query Match 93.8%; Score 30; DB 11; Length 108;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7

Db 80 GFYRN 84
|||||

RESULT 14

US-10-156-761-11215
; Sequence 11215, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HAITORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11215
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11215

Query Match 93.8%; Score 30; DB 15; Length 113;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7

Db 51 GFYRN 55
|||||

RESULT 15

US-09-764-877-1934
; Sequence 1934, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1934
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-877-1934

Query Match 93.8%; Score 30; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7

Db 62 GFYRN 66
|||||

Search completed: December 8, 2003, 09:34:03
Job time : 6.22222 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:31 ; Search time 3.25253 Seconds
(without alignments)
91.060 Million cell updates/sec

Title: US-09-498-556C-79
Perfect score: 32
Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	5	2 US-08-465-380-70	Sequence 70, Appl
2	30	93.8	5	2 US-08-486-397-70	Sequence 70, Appl
3	30	93.8	5	2 US-08-486-399-70	Sequence 70, Appl
4	30	93.8	5	2 US-08-461-965-70	Sequence 70, Appl
5	30	93.8	5	2 US-08-634-641-70	Sequence 70, Appl
6	30	93.8	5	3 US-09-249-471-70	Sequence 70, Appl
7	30	93.8	5	3 US-09-249-472-70	Sequence 70, Appl
8	30	93.8	5	3 US-09-249-451-70	Sequence 70, Appl
9	30	93.8	5	3 US-08-809-455-70	Sequence 70, Appl
10	30	93.8	5	3 US-09-249-461-70	Sequence 70, Appl
11	30	93.8	5	3 US-09-249-448-70	Sequence 70, Appl
12	30	93.8	5	4 US-09-249-473-70	Sequence 70, Appl
13	30	93.8	7	2 US-08-465-380-79	Sequence 79, Appl
14	30	93.8	7	2 US-08-486-397-79	Sequence 79, Appl
15	30	93.8	7	2 US-08-486-399-79	Sequence 79, Appl
16	30	93.8	7	2 US-08-461-965-79	Sequence 79, Appl
17	30	93.8	7	2 US-08-634-641-79	Sequence 79, Appl
18	30	93.8	7	3 US-09-249-471-79	Sequence 79, Appl
19	30	93.8	7	3 US-09-249-472-79	Sequence 79, Appl
20	30	93.8	7	3 US-09-249-451-79	Sequence 79, Appl
21	30	93.8	7	3 US-08-809-455-79	Sequence 79, Appl
22	30	93.8	7	3 US-09-249-461-79	Sequence 79, Appl
23	30	93.8	7	3 US-09-249-448-79	Sequence 79, Appl
24	30	93.8	7	4 US-09-249-473-79	Sequence 79, Appl
25	30	93.8	78	2 US-08-465-380-47	Sequence 47, Appl
26	30	93.8	78	2 US-08-486-397-47	Sequence 47, Appl
27	30	93.8	78	2 US-08-486-399-47	Sequence 47, Appl

28	30	93.8	78	2 US-08-461-965-47	Sequence 47, Appl
29	30	93.8	78	2 US-08-634-641-47	Sequence 47, Appl
30	30	93.8	78	3 US-09-249-471-47	Sequence 47, Appl
31	30	93.8	78	3 US-09-249-472-47	Sequence 47, Appl
32	30	93.8	78	3 US-09-249-451-47	Sequence 47, Appl
33	30	93.8	78	3 US-08-809-455-47	Sequence 47, Appl
34	30	93.8	78	3 US-09-249-461-47	Sequence 47, Appl
35	30	93.8	78	3 US-09-249-448-47	Sequence 47, Appl
36	30	93.8	78	4 US-09-249-473-47	Sequence 47, Appl
37	30	93.8	82	2 US-08-465-380-25	Sequence 25, Appl
38	30	93.8	82	2 US-08-486-397-49	Sequence 49, Appl
39	30	93.8	82	2 US-08-480-478-54	Sequence 54, Appl
40	30	93.8	82	2 US-08-486-397-25	Sequence 25, Appl
41	30	93.8	82	2 US-08-486-397-49	Sequence 49, Appl
42	30	93.8	82	2 US-08-486-399-25	Sequence 25, Appl
43	30	93.8	82	2 US-08-486-399-49	Sequence 49, Appl
44	30	93.8	82	2 US-08-461-965-25	Sequence 25, Appl
45	30	93.8	82	2 US-08-461-965-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-08-465-380-70
; Sequence 70, Application US/08465380
; Patent No. 5863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemaans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATOIDE-EXTRACTED ANTICONGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,380
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal fragment
US-08-465-380-70

Query Match 93.8%; Score 30; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
 Db 1 GFYRN 5

RESULT 2
 US-08-486-397-70
 ; Sequence 70, Application US/08486397
 ; Patent No. 5866542
 ; GENERAL INFORMATION:
 ; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
 ; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
 ; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
 ; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
 ; APPLICANT: Peter W. Bergum
 ; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
 ; NUMBER OF SEQUENCES: 357
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: 633 West Fifth Street
 ; STREET: Suite 4700
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 90071

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/486,397
 FILING DATE: June 5, 1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/326,110
 FILING DATE: October 18, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BIGGS, SUZANNE L.
 REGISTRATION NUMBER: 30,158
 REFERENCE/DOCKET NUMBER: 213/269
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 70:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal fragment
 US-08-486-397-70

Query Match 93.8%; Score 30; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
 Db 1 GFYRN 5

RESULT 3
 US-08-486-399-70
 ; Sequence 70, Application US/08486399

Patent No. 5866543
 ; GENERAL INFORMATION:
 ; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
 ; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
 ; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
 ; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
 ; APPLICANT: Peter W. Bergum
 ; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
 ; NUMBER OF SEQUENCES: 356
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: 633 West Fifth Street
 ; STREET: Suite 4700
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 90071

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/486,399
 FILING DATE: June 5, 1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/326,110
 FILING DATE: October 18, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BIGGS, SUZANNE L.
 REGISTRATION NUMBER: 30,158
 REFERENCE/DOCKET NUMBER: 213/270
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 70:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal fragment
 US-08-486-399-70

Query Match 93.8%; Score 30; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
 Db 1 GFYRN 5

RESULT 4
 US-08-461-965-70
 ; Sequence 70, Application US/08461965
 ; Patent No. 5872098
 ; GENERAL INFORMATION:
 ; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
 ; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
 ; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
 ; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
 ; APPLICANT: Peter W. Bergum
 ; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
 ; NUMBER OF SEQUENCES: 356
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: 633 West Fifth Street

STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,965
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 210/243
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
US-08-461-965-70

Query Match 93.8%; Score 30; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 1 GFYRN 5

RESULT 5
US-08-634-641-70
Sequence 70, Application US/08634641
Patent No. 5955294
GENERAL INFORMATION:
APPLICANT: Vlasuk, George P. Vlasuk
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Mensens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Gansmans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
APPLICANT: Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/634,641
FILING DATE: April 19, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 219/136
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
US-08-634-641-70

Query Match 93.8%; Score 30; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 1 GFYRN 5

RESULT 6
US-09-249-471-70
Sequence 70, Application US/09249471
Patent No. 6040441
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Mensens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Gansmans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
APPLICANT: Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,471
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/809,455
FILING DATE: April 17, 1997
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
US-09-249-471-70

Query Match 93.8%; Score 30; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

Qy 3 GFYRN 7
Db 1 GFYRN 5

RESULT 7

US-09-249-472-70
Sequence 70, Application US/09249472
Patent No. 6046318
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Ganssemans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
APPLICANT: Berghum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
ZIP: U.S.A.
90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,472
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/809,455
FILING DATE: April 17, 1997
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
US-09-249-472-70

Query Match 93.8%; Score 30; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

Qy 3 GFYRN 7
Db 1 GFYRN 5

RESULT 8

US-09-249-451-70
Sequence 70, Application US/09249451
Patent No. 6087487
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Ganssemans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
APPLICANT: Berghum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STATE: Los Angeles
COUNTRY: California
ZIP: U.S.A.
90071

STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,451
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/809,455
FILING DATE: April 17, 1997
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
US-09-249-451-70

Query Match 93.8%; Score 30; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPYRN 7
Db 1 GPYRN 5

RESULT 9
US-08-809-455-70
Sequence 70, Application US/08809455
Patent No. 6090916
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Ganssemans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
TITLE OF INVENTION: NEMATOIDE-EXTRACTED SERINE PROTEASE
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,455
FILING DATE: April 17, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
US-08-809-455-70

Query Match 93.8%; Score 30; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GPYRN 7
Db 1 GPYRN 5

RESULT 10
US-09-249-461-70
Sequence 70, Application US/09249461
Patent No. 6096877
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Ganssemans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
TITLE OF INVENTION: NEMATOIDE-EXTRACTED SERINE PROTEASE
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,461
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/809,455
FILING DATE: April 17, 1997
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
US-09-249-461-70

Query Match 93.8%; Score 30; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 1 GFYRN 5

RESULT 11

US-09-249-448-70
Sequence 70, Application US/09249448
Patent No. 6121435
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Larocche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Ganssemans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
TITLE OF INVENTION: NEMATOIDE-EXTRACTED SERINE PROTEASE
INHIBITORS AND ANTICOAGULANT

TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,448
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/809,455
FILING DATE: April 17, 1997
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal fragment
US-09-249-448-70

Query Match 93.8%; Score 30; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 1 GFYRN 5

RESULT 12

US-09-249-473-70
Sequence 70, Application US/09249473
Patent No. 6534629
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
APPLICANT: Stanssens, Patrick Eric Hugo
APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Larocche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Ganssemans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew

APPLICANT: Bergum, Peter W.
 TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
 TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
 NUMBER OF SEQUENCES: 356
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/249,473
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/809,455
 FILING DATE: April 17, 1997
 APPLICATION NUMBER: PCT/US95/13231
 FILING DATE: October 17, 1995
 APPLICATION NUMBER: 08/486,399
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/486,397
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/465,380
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/461,965
 FILING DATE: June 5, 1995
 APPLICATION NUMBER: 08/326,110
 FILING DATE: October 18, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BIGGS, SUZANNE L.
 REGISTRATION NUMBER: 30,158
 REFERENCE/DOCKET NUMBER: 216/270
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 70:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal fragment

Query Match 93.8%; Score 30; DB 4; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
 Db 1 GFYRN 5

RESULT 13
 US-08-465-380-79
 Sequence 79, Application US/08465380
 Patent No. 5863894
 GENERAL INFORMATION:
 APPLICANT: George P. Vlasuk, Patric H. Stanssens,
 APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
 APPLICANT: Yves R. Laroche, Laurent S. Jespers,
 APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
 APPLICANT: Peter W. Bergum

TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
 TITLE OF INVENTION: PROTEIN
 NUMBER OF SEQUENCES: 356
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Word Perfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/465,380
 FILING DATE: June 5, 1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/326,110
 FILING DATE: October 18, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BIGGS, SUZANNE L.
 REGISTRATION NUMBER: 30,158
 REFERENCE/DOCKET NUMBER: 213/268
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 79:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal fragment
 FEATURE:
 OTHER INFORMATION: Xaa in locations 1 and 2
 OTHER INFORMATION: is an amino acid, provided
 OTHER INFORMATION: that at least one Xaa is Glu or
 OTHER INFORMATION: Asp.

US-08-465-380-79
 Query Match 93.8%; Score 30; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.5e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
 Db 3 GFYRN 7

RESULT 14
 US-08-486-397-79
 Sequence 79, Application US/08486397
 Patent No. 5866542
 GENERAL INFORMATION:
 APPLICANT: George P. Vlasuk, Patric H. Stanssens,
 APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
 APPLICANT: Yves R. Laroche, Laurent S. Jespers,
 APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
 APPLICANT: Peter W. Bergum
 TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
 NUMBER OF SEQUENCES: 357
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Fifth Street
 CITY: Los Angeles

```
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,397
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal fragment
; FEATURE:
; OTHER INFORMATION: Xaa in locations 1 and 2
; OTHER INFORMATION: is an amino acid, provided
; OTHER INFORMATION: that at least one Xaa is Glu or
; OTHER INFORMATION: Asp.
; US-08-486-399-79

Query Match 93.8%; Score 30; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 3 GFYRN 7

Search completed: December 8, 2003, 09:20:29
Job time: 4.25253 secs

; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,397
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/269
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal fragment
; FEATURE:
; OTHER INFORMATION: Xaa in locations 1 and 2
; OTHER INFORMATION: is an amino acid, provided
; OTHER INFORMATION: that at least one Xaa is Glu or
; OTHER INFORMATION: Asp.
; US-08-486-397-79

Query Match 93.8%; Score 30; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 3 GFYRN 7

RESULT 15
US-08-486-399-79
; Sequence 79, Application US/08486399
; Patent No. 5866543
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Ganssens, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 9.54545 Seconds
(without alignments)
116.399 Million cell updates/sec

Title: US-09-498-556C-79
Perfect score: 32
Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	5	17	AA1991721
2	30	93.8	5	20	AA1930440
3	30	93.8	5	21	AA1915325
4	30	93.8	7	17	AA191730
5	30	93.8	7	20	AA1930449
6	30	93.8	7	21	AA1915334
7	30	93.8	78	20	AA1930420
8	30	93.8	78	21	AA1915305
9	30	93.8	82	20	AA1930399

10	30	93.8	82	20	AA1930422	Mature nematode ex
11	30	93.8	82	21	AA1915293	A. ceylanicum nema
12	30	93.8	82	21	AA1915307	A. ceylanicum nema
13	30	93.8	83	20	AA1930425	Mature nematode ex
14	30	93.8	84	20	AA1930423	Mature nematode ex
15	30	93.8	84	20	AA1930424	Mature nematode ex
16	30	93.8	84	20	AA1930432	Mature nematode ex
17	30	93.8	84	21	AA1915317	A. caninum nematod
18	30	93.8	86	22	AA191998	Human digestive sy
19	30	93.8	89	20	AA1930421	Mature nematode ex
20	30	93.8	91	17	AA191701	AcanaPc2. Ancylos
21	30	93.8	91	20	AA1930393	Nematode extracted
22	30	93.8	91	20	AA1930454	Nematode extracted
23	30	93.8	91	21	AA1915346	A. caninum nematod
24	30	93.8	97	17	AA191709	AcanaPc1. Ancylos
25	30	93.8	97	20	AA1930407	Nematode extracted
26	30	93.8	97	22	AA1914986	Human nervous syst
27	30	93.8	102	20	AA1930401	Nematode extracted
28	30	93.8	102	21	AA1915295	A. duodenale nemat
29	30	93.8	108	20	AA1930398	Nematode extracted
30	30	93.8	108	21	AA1915292	A. ceylanicum nema
31	30	93.8	124	21	AA1970530	Maize plastid targ
32	30	93.8	138	22	AA1903987	Human musculoskele
33	30	93.8	138	24	AA1913281	Novel human muscul
34	30	93.8	161	20	AA1930438	Mature nematode ex
35	30	93.8	162	20	AA1930436	Mature nematode ex
36	30	93.8	162	20	AA1930437	Mature nematode ex
37	30	93.8	162	21	AA1915321	A. caninum nematod
38	30	93.8	162	21	AA1915322	A. caninum nematod
39	30	93.8	171	20	AA1930435	Mature nematode ex
40	30	93.8	180	17	AA191713	AduNAP7. Ancylost
41	30	93.8	180	20	AA1930388	Nematode extracted
42	30	93.8	181	17	AA191711	AcanaP45. Ancylos
43	30	93.8	181	17	AA191712	AcanaP47. Ancylos
44	30	93.8	181	20	AA1930409	Nematode extracted
45	30	93.8	181	20	AA1930410	Nematode extracted

ALIGNMENTS

RESULT 1
AA191721
ID AA191721 standard; Protein; 5 AA.
XX
AC AA191721;
XX
DT 25-MAR-2003 (updated)
DT 17-NOV-1996 (first entry)
XX
DE NAP subsequence.
XX
KW AcanAP; HpoNAP; NamNAP; AcanAP; AduNAP; anticoagulant;
KW nematode-extracted anticoagulant protein; serine protease;
KW nematode; thrombosis; parasitic worm.
XX
OS Synthetic.
XX
PN WO9612021-A2.
XX
PD 25-APR-1996.
XX
PF 17-OCT-1995; 95WO-US13231.
XX
PR 05-JUN-1995; 95US-0486399.
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
XX
PA (CORV-) CORVAS INT INC.
XX
PI Vlaauk GP, Stanssens PEH, Messens JHL, Lauwereys MJ, Laroche YR;

PI Jespers LS, Gansemans YGJ, Moyle M, Bergum PW;
 XX WPT; 1996-222007/22.
 XX
 XX Proteins with anticoagulant and/or serine protease inhibitory
 PT activity - isolated from nematodes and useful to inhibit blood
 PT coagulation
 XX
 XX Claim 10; Page 144; 243pp; English.
 XX
 CC Proteins with anticoagulant and/or serine protease inhibitory
 CC activity, isolated from nematodes, are useful to inhibit blood
 CC coagulation. The proteins can be added to blood collection tubes
 CC defining the collection of mammalian plasma. They are also useful
 CC to prevent or inhibit thrombosis, and may be given alone or in
 CC combination with other therapeutic or in vivo diagnostic agents.
 CC The proteins can serve as immunogens to raise antibodies for use in
 CC the diagnosis and identification of NAP concn. levels in biological
 CC fluids, e.g. to detect mammalian infection with a parasitic worm.
 CC They can also be used as immunogens in prophylactic and therapeutic
 CC vaccines against parasitic worm infection. The proteins may
 CC double the clotting time of human plasma in prothrombin time assays
 CC when present at 10-50 nMol, and double the clotting time of human
 CC plasma in activated partial thrombin time assays when present
 CC at 10-100 nMol.
 CC The anticoagulant proteins are pref. derived from
 CC Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator
 CC americanus or Heligmosomoides polygyrus.
 CC The proteins pref. have 2 NAP domains and specifically inhibit
 CC the catalytic activity of the factor VIIa/TF complex in the
 CC presence of factor Xa or a catalytically inactive factor Xa deriv.,
 CC do not specifically inhibit the activation of factor VIIa in the
 CC absence of TF and do not specifically inhibit prothrombinase.
 CC Proteins given in AAR91720-R91732 are preferred subsequences
 CC of a generic NAP sequence.
 CC (Updated on 25-NAR-2003 to correct PI field.)
 XX
 XX Sequence 5 AA;

Query Match 93.8%; Score 30; DB 17; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
 |||||
 Db 1 GFYRN 5

RESULT 2
 AAY30440
 ID AAY30440 standard; Peptide; 5 AA.
 XX
 AC AAY30440;
 XX
 DT 15-NOV-1999 (first entry)
 XX
 DE Nematode extracted anticoagulant protein fragment.
 XX
 KW Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX
 OS Unidentified.
 XX
 FN US5955294-A.
 XX
 PD 21-SEP-1999.
 XX
 PF 19-APR-1996; 96US-0634641.
 XX
 PR 19-APR-1996; 96US-0634641.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.
 PR 05-JUN-1995; 95US-0486399.
 PR 17-OCT-1995; 95WO-US13231.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
 PI Vlasuk GP;
 XX
 DR WPI; 1999-539569/45.
 XX
 PT Screening an isolated protein for Nematode-extracted Anticoagulant
 PT Protein domains
 XX
 PS Disclosure; Column 151; 197pp; English.
 XX
 CC The present sequence represents a nematode extracted anticoagulant
 CC protein (NAP) fragment. The protein has activity as an anticoagulant
 CC and/or serine protease inhibitor. The protein contains at least one
 CC NAP domain which has selective inhibitory activity for factor VIIa/TF.
 CC The specification describes a method for screening an isolated protein
 CC at least one domain for factor VIIa/TF selective inhibitory activity.
 CC The method comprises determining the time to clotting effected by a
 CC concentration of the isolated protein in an ex vivo prothrombin time
 CC (PT) assay and an ex vivo activated partial thromboplastin time (aPTT)
 CC assay; calculating prolongation of clotting effected by the isolated
 CC protein in each of the PT and aPTT assay, with respect to a baseline
 CC clotting value for each assay, where prolongation of clotting is
 CC calculated as fold elevation of clotting time relative to a baseline
 CC clotting value, where a doubling of clotting time is deemed a two-fold
 CC elevation; and calculating a PT to aPTT prolongation ratio, where a
 CC ratio at least one is indicative of factor VIIa/TF inhibitory activity.
 CC The method is useful for determining if a protein has factor VIIa/TF
 CC inhibitory activity.
 XX
 XX Sequence 5 AA;

Query Match 93.8%; Score 30; DB 20; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
 |||||
 Db 1 GFYRN 5

RESULT 3
 AAB15325
 ID AAB15325 standard; Peptide; 5 AA.
 XX
 AC AAB15325;
 XX
 DT 19-DEC-2000 (first entry)
 XX
 DE NAP domain fragment #2.
 XX
 KW Nematode-extracted anticoagulant protein; NAP domain; blood clotting;
 KW canine hookworm; thrombosis; vaccine.
 XX
 OS Unidentified.
 XX
 FN US6087487-A.
 XX
 PD 11-JUL-2000.
 XX
 PF 12-FEB-1999; 99US-0249451.
 XX
 PR 17-OCT-1995; 95WO-US13231.
 PR 17-APR-1997; 97US-0809455.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.

PR 05-JUN-1995; 95US-0486397.
 XX 05-JUN-1995; 95US-0486399.
 PA (CORV-) CORVAS INT INC.
 XX Lauwereys MJ, Stanssens PEH, Jespers LS, Ganssemans YGJ, Moyle M;
 PI Bergum FW, Messens JHL, Laroche YR, Vlasuk GP;
 XX WPI; 2000-531359/48.
 XX
 XX New cDNA molecule encoding a protein having factor Xa inhibitory
 PT activity for preventing and treating blood clotting disorders,
 PT comprises nematode-extracted anticoagulant protein domains -
 XX
 XX Claim 3; Column 285; 197pp; English.
 XX
 XX The present sequence is a fragment of the NAP domain (see AAB15347),
 CC which is found in all nematode-extracted anticoagulant proteins (NAPs).
 CC Proteins of this kind have been shown to be effective at preventing
 CC blood clotting without causing excessive bleeding. They can be used in
 CC blood collection tubes to aid the isolation of plasma from the blood, to
 CC prevent thrombosis which may be linked to the rupture of an
 CC atherosclerotic plaque, acute myocardial infarction, angina,
 CC thrombolytic therapy, percutaneous transluminal coronary angioplasty,
 CC disseminated intravascular coagulopathy, infection, cancer and septic
 CC shock, and to produce antibodies. In the latter instance, the antibodies
 CC can be raised in order to detect infection by nematodes (the coding
 CC sequences can also be used for this) or as diagnostic tests. The proteins
 CC can also be used as vaccines against nematode parasites.
 XX
 XX Sequence 5 AA;
 SQ
 Query Match 93.8%; Score 30; DB 21; Length 5;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GFYRN 7
 Db 1 GFYRN 5
 |||||
 |||||
 RESULT 4
 ID AAR91730 standard; Protein; 7 AA.
 AC AAR91730;
 XX
 XX 25-MAR-2003 (updated)
 DT 17-NOV-1996 (first entry)
 XX
 XX NAP subsequence.
 XX
 XX AcanAP; HponAP; NamNAP; AceNAP; AduNAP; anticoagulant;
 KW nematode-extracted anticoagulant protein; serine protease;
 KW nematode; thrombosis; parasitic worm.
 XX
 XX Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 1
 FT /label= Glu, Asp, OTHER
 FT /note= "at least one of residue 1 or residue 2
 FT is Glu or Asp"
 FT
 FT Misc-difference 2
 FT /label= Glu, Asp, OTHER
 FT /note= "at least one of residue 1 or residue 2
 FT is Glu or Asp"
 FT
 XX WO9612021-A2.
 XX
 XX 25-APR-1996.
 PD
 XX 17-OCT-1995; 95WO-US13231.
 PF

XX 05-JUN-1995; 95US-0486399.
 PR 18-OCT-1994; 94US-0326110.
 XX 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0463380.
 PR 05-JUN-1995; 95US-0486397.
 XX
 XX (CORV-) CORVAS INT INC.
 XX
 XX Vlasuk GP, Stanssens PEH, Messens JHL, Lauwereys MJ, Laroche YR;
 PI Jespers LS, Ganssemans YGJ, Moyle M, Bergum FW;
 XX WPI; 1996-222007/22.
 DR
 XX
 XX Proteins with anticoagulant and/or serine protease inhibitory
 PT activity - isolated from nematodes and useful to inhibit blood
 PT coagulation
 XX
 XX Claim 26; Page 147; 243pp; English.
 XX
 XX Proteins with anticoagulant and/or serine protease inhibitory
 CC activity, isolated from nematodes, are useful to inhibit blood
 CC coagulation. The proteins can be added to blood collection tubes
 CC defining the collection of mammalian plasma. They are also useful
 CC to prevent or inhibit thrombosis, and may be given alone or in
 CC combination with other therapeutic or in vivo diagnostic agents.
 CC The proteins can serve as immunogens to raise antibodies for use in
 CC the diagnosis and identification of NAP concn. levels in biological
 CC fluids, e.g. to detect mammalian infection with a parasitic worm.
 CC They can also be used as immunogens in prophylactic and therapeutic
 CC vaccines against parasitic worm infection. The proteins may
 CC double the clotting time of human plasma in prothrombin time assays
 CC when present at 10-50 nMol, and double the clotting time of human
 CC plasma in activated partial thrombin time assays when present
 CC at 10-100 nMol.
 CC The anticoagulant proteins are pref. derived from
 CC Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator
 CC americanus or Heligmosomoides polygyrus.
 CC The proteins pref. have 2 NAP domains and specifically inhibit
 CC the catalytic activity of the factor VIIa/TF complex in the
 CC presence of factor Xa or a catalytically inactive factor Xa deriv.,
 CC do not specifically inhibit the activation of factor VIII in the
 CC absence of TF and do not specifically inhibit prothrombinase.
 CC Proteins given in AAR91720-R91732 are preferred subsequences
 CC of a generic NAP sequence.
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 XX Sequence 7 AA;
 SQ
 Query Match 93.8%; Score 30; DB 17; Length 7;
 Best Local Similarity 100.0%; Pred. No. 9.3e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GFYRN 7
 Db 3 GFYRN 7
 |||||
 |||||
 RESULT 5
 ID AAY30449 standard; Peptide; 7 AA.
 XX
 XX AAY30449;
 AC
 XX
 XX 15-NOV-1999 (first entry)
 DT
 XX Nematode extracted anticoagulant protein fragment.
 DE
 XX Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX
 XX Unidentified.
 OS
 XX

XX PD 21-SEP-1999.
 XX PF 19-APR-1996; 96US-0634641.
 XX PR 19-APR-1996; 96US-0634641.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.
 PR 05-JUN-1995; 95US-0486397.
 PR 05-JUN-1995; 95US-0486399.
 PR 17-OCT-1995; 95WO-US13231.
 XX (CORV-) CORVAS INT INC.
 XX PA Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
 XX PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
 PI Vlasuk GP;
 XX WPI; 1999-539569/45.
 XX Screening an isolated protein for Nematode-extracted Anticoagulant
 PT Protein domains
 PT Disclosure; Columns 133-134; 197pp; English.
 XX The present sequence represents a nematode extracted anticoagulant
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine
 CC protease inhibitor. The protein contains at least one NAP domain which
 CC has selective inhibitory activity for factor VIIa/TF. The specification
 CC describes a method for screening an isolated protein at least one domain
 CC for factor VIIa/TF selective inhibitory activity. The method comprises
 CC determining the time to clotting effected by a concentration of the
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
 CC activated partial thromboplastin time (APTT) assay; calculating
 CC prolongation of clotting effected by the isolated protein in each of
 CC the PT and APTT assay, with respect to a baseline clotting value for
 CC each assay, where prolongation of clotting is calculated as fold
 CC elevation of clotting time relative to a baseline clotting value, where
 CC a doubling of clotting time is deemed a two-fold elevation; and
 CC calculating a PT to APTT prolongation ratio, where a ratio at least
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is
 CC useful for determining if a protein has factor VIIa/TF inhibitory
 CC activity.
 XX SQ Sequence 78 AA;
 Query Match 93.8%; Score 30; DB 20; Length 78;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GFYRN 7
 Db 54 GFYRN 58
 RESULT 8
 AAB15305
 ID AAB15305 standard; Protein; 78 AA.
 XX AC AAB15305;
 XX DT 19-DEC-2000 (first entry)
 XX DE A. caninum nematode-extracted anticoagulant protein AcanAP31,42,46.
 XX KW Nematode-extracted anticoagulant protein; AcanAP31; AcanAP42; AcanAP46;
 KW canine hookworm; blood clotting; thrombosis; vaccine.
 XX OS Ancylostoma caninum.
 XX PF US6087487-A.
 XX PR

PD 11-JUL-2000.
 XX 12-FEB-1999; 99US-0249451.
 XX PR 17-OCT-1995; 95WO-US13231.
 PR 17-APR-1997; 97US-0809455.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.
 PR 05-JUN-1995; 95US-0486397.
 PR 05-JUN-1995; 95US-0486399.
 XX (CORV-) CORVAS INT INC.
 XX PA Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;
 XX PI Bergum PW, Messens JHL, Laroche YR, Vlasuk GP;
 XX WPI; 2000-531359/48.
 XX N-PSDB; AAA73378.
 XX New cDNA molecule encoding a protein having factor Xa inhibitory
 PT activity for preventing and treating blood clotting disorders,
 PT comprises nematode-extracted anticoagulant protein domains -
 XX Disclosure; Fig 16; 197pp; English.
 XX The present sequence comprises the Ancylostoma caninum
 CC nematode-extracted anticoagulant proteins AcanAP31, AcanAP42 and
 CC AcanAP46. Proteins of this kind have been shown to be effective at
 CC preventing blood clotting without causing excessive bleeding. The
 CC proteins can be used in blood collection tubes to aid the isolation of
 CC plasma from the blood, to prevent thrombosis which may be linked to the
 CC rupture of an atherosclerotic plaque, acute myocardial infarction,
 CC angina, thrombolytic therapy, percutaneous transluminal coronary
 CC angioplasty, disseminated intravascular coagulopathy, infection, cancer
 CC and septic shock, and to produce antibodies. In the latter instance, the
 CC antibodies can be raised in order to detect infection by nematodes (the
 CC coding sequence can also be used for this) or as diagnostic tests. The
 CC proteins can also be used as a vaccine against nematode parasites.
 XX SQ Sequence 78 AA;
 Query Match 93.8%; Score 30; DB 21; Length 78;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GFYRN 7
 Db 54 GFYRN 58
 RESULT 9
 AAY30399
 ID AAY30399 standard; Protein; 82 AA.
 XX AC AAY30399;
 XX DT 15-NOV-1999 (first entry)
 XX DE Nematode extracted anticoagulant protein AcanAP4d2.
 XX KW Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX OS Ancylostoma ceylanicum.
 XX PN US5955294-A.
 XX PD 21-SEP-1999.
 XX PF 19-APR-1996; 96US-0634641.
 XX PR 19-APR-1996; 96US-0634641.

```
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
PR 05-JUN-1995; 95US-0486399.
PR 17-OCT-1995; 95WO-US13231.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
PI Vlasuk GP;
XX
XX WPI; 1999-539569/45.
XX
XX Screening an isolated protein for Nematode-extracted Anticoagulant
PT Protein domains
XX
XX Disclosure; Column 111-112; 197pp; English.
XX
XX The present sequence represents a nematode extracted anticoagulant
CC protein (NAP). The protein has activity as an anticoagulant and/or serine
CC protease inhibitor. The protein contains at least one NAP domain which
CC has selective inhibitory activity for factor VIIa/TF. The specification
CC describes a method for screening an isolated protein at least one domain
CC for factor VIIa/TF selective inhibitory activity. The method comprises
CC determining the time to clotting effected by a concentration of the
CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
CC activated partial thromboplastin time (APTT) assay; calculating
CC prolongation of clotting effected by the isolated protein in each of
CC the PT and APTT assay, with respect to a baseline clotting value for
CC each assay, where prolongation of clotting is calculated as fold
CC elevation of clotting time relative to a baseline clotting value, where
CC a doubling of clotting time is deemed a two-fold elevation; and
CC calculating a PT to APTT prolongation ratio, where a ratio at least
CC one is indicative of factor VIIa/TF inhibitory activity. The method is
CC useful for determining if a protein has factor VIIa/TF inhibitory
CC activity.
XX
SQ Sequence 82 AA;

Query Match 93.8%; Score 30; DB 20; Length 82;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 51 GFYRN 55
|||||

RESULT 10
AAAY30422
ID AAAY30422 standard; Protein; 82 AA.
XX
XX AAAY30422;
XX
XX 15-NOV-1999 (first entry)
XX
XX Mature nematode extracted anticoagulant protein AceNAP4d2.
DE
XX Nematode extracted anticoagulant protein; NAP; anticoagulant;
KW serine protease inhibitor; NAP domain; factor VIIa/TF.
KW
XX
XX Ancylostoma ceylanicum.
OS
XX US5955294-A.
PN
XX
XX 21-SEP-1999.
PD
XX
XX 19-APR-1996; 96US-0634641.
XX
XX 19-APR-1996; 96US-0634641.
PR
PR 18-OCT-1994; 94US-0326110.
PR
```

```
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
PR 17-OCT-1995; 95WO-US13231.
XX
XX (CORV-) CORVAS INT INC.
XX
XX Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
PI Vlasuk GP;
XX
XX WPI; 1999-539569/45.
XX
XX Screening an isolated protein for Nematode-extracted Anticoagulant
PT Protein domains
XX
XX Disclosure; Columns 135-136; 197pp; English.
XX
XX The present sequence represents a nematode extracted anticoagulant
CC protein (NAP). The protein has activity as an anticoagulant and/or serine
CC protease inhibitor. The protein contains at least one NAP domain which
CC has selective inhibitory activity for factor VIIa/TF. The specification
CC describes a method for screening an isolated protein at least one domain
CC for factor VIIa/TF selective inhibitory activity. The method comprises
CC determining the time to clotting effected by a concentration of the
CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
CC activated partial thromboplastin time (APTT) assay; calculating
CC prolongation of clotting effected by the isolated protein in each of
CC the PT and APTT assay, with respect to a baseline clotting value for
CC each assay, where prolongation of clotting is calculated as fold
CC elevation of clotting time relative to a baseline clotting value, where
CC a doubling of clotting time is deemed a two-fold elevation; and
CC calculating a PT to APTT prolongation ratio, where a ratio at least
CC one is indicative of factor VIIa/TF inhibitory activity. The method is
CC useful for determining if a protein has factor VIIa/TF inhibitory
CC activity.
XX
SQ Sequence 82 AA;

Query Match 93.8%; Score 30; DB 20; Length 82;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 51 GFYRN 55
|||||

RESULT 11
AAB15293
ID AAB15293 standard; Protein; 82 AA.
XX
XX AAB15293;
XX
XX 19-DEC-2000 (first entry)
XX
XX A. ceylanicum nematode-extracted anticoagulant protein AceNAP4d2 #1.
DE
XX Nematode-extracted anticoagulant protein; AceNAP4d2; blood clotting;
KW canine hookworm; thrombosis; vaccine.
KW
XX Ancylostoma ceylanicum.
OS
XX US6087487-A.
PN
XX
XX 11-JUL-2000.
PD
XX
XX 12-FEB-1999; 99US-0249451.
XX
XX 17-OCT-1995; 95WO-US13231.
XX
XX 17-APR-1997; 97US-0809455.
PR
PR 18-OCT-1994; 94US-0326110.
PR
```


PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.
 PR 05-JUN-1995; 95US-0486397.
 PR 05-JUN-1995; 95US-0486399.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;
 PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;
 XX
 DR WPI; 2000-531359/48.
 XX
 PT New cDNA molecule encoding a protein having factor Xa inhibitory
 PT activity for preventing and treating blood clotting disorders,
 PT comprises nematode-extracted anticoagulant protein domains -
 XX
 PS Disclosure; Fig 11; 197pp; English.
 XX
 CC The present sequence is the Ancylostoma ceylanicum nematode-extracted
 CC anticoagulant protein AcenAP4d2. Proteins of this kind have been shown to
 CC be effective at preventing blood clotting without causing excessive
 CC bleeding. The protein can be used in blood collection tubes to aid the
 CC isolation of plasma from the blood, to prevent thrombosis which may be
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,
 CC cancer and septic shock, and to produce antibodies. In the latter
 CC instance, the antibodies can be raised in order to detect infection by
 CC nematodes or as diagnostic tests. The protein can also be used as a
 CC vaccine against nematode parasites.
 XX
 SQ Sequence 82 AA;
 Query Match 93.8%; Score 30; DB 21; Length 82;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3 GFYRN 7
 Db 51 GFYRN 55
 |||||
 RESULT 12
 AABI5307
 ID AABI5307 standard; Protein; 82 AA.
 AC AABI5307;
 XX
 DT 19-DEC-2000 (first entry)
 XX
 DE A. ceylanicum nematode-extracted anticoagulant protein AcenAP4d2 #2.
 XX
 KW Nematode-extracted anticoagulant protein; AcenAP4d2; blood clotting;
 KW canine hookworm; thrombosis; vaccine.
 XX
 OS Ancylostoma ceylanicum.
 XX
 PN US6087487-A.
 XX
 PD 11-JUL-2000.
 XX
 PF 12-FEB-1999; 99US-0249451.
 XX
 PR 17-OCT-1995; 95WO-US13231.
 PR 17-APR-1997; 97US-0809455.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.
 PR 05-JUN-1995; 95US-0486397.
 PR 05-JUN-1995; 95US-0486399.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;
 PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;
 XX
 DR WPI; 2000-531359/48.
 XX
 PT New cDNA molecule encoding a protein having factor Xa inhibitory
 PT activity for preventing and treating blood clotting disorders,
 PT comprises nematode-extracted anticoagulant protein domains -
 XX
 PS Disclosure; Fig 11; 197pp; English.
 XX
 CC The present sequence is the Ancylostoma ceylanicum nematode-extracted
 CC anticoagulant protein AcenAP4d2. Proteins of this kind have been shown to
 CC be effective at preventing blood clotting without causing excessive
 CC bleeding. The protein can be used in blood collection tubes to aid the
 CC isolation of plasma from the blood, to prevent thrombosis which may be
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,
 CC cancer and septic shock, and to produce antibodies. In the latter
 CC instance, the antibodies can be raised in order to detect infection by
 CC nematodes or as diagnostic tests. The protein can also be used as a
 CC vaccine against nematode parasites.
 XX
 SQ Sequence 82 AA;
 Query Match 93.8%; Score 30; DB 21; Length 82;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3 GFYRN 7
 Db 51 GFYRN 55
 |||||
 RESULT 13
 AAY30425
 ID AAY30425 standard; Protein; 83 AA.
 XX
 AC AAY30425;
 XX
 DT 15-NOV-1999 (first entry)
 XX
 DE Mature nematode extracted anticoagulant protein AduNAP7di.
 XX
 KW Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX
 OS Ancylostoma duodenale.
 XX
 PN US5955294-A.
 XX
 PD 21-SEP-1999.
 XX
 PF 19-APR-1996; 96US-0634641.
 XX
 PR 19-APR-1996; 96US-0634641.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.
 PR 05-JUN-1995; 95US-0486397.
 PR 05-JUN-1995; 95US-0486399.
 PR 17-OCT-1995; 95WO-US13231.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
 PI Vlasuk GP;
 XX
 DR WPI; 1999-539569/45.
 XX
 PT Screening an isolated protein for Nematode-extracted Anticoagulant

PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;
 PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;
 XX
 DR WPI; 2000-531359/48.
 XX
 PT New cDNA molecule encoding a protein having factor Xa inhibitory
 PT activity for preventing and treating blood clotting disorders,
 PT comprises nematode-extracted anticoagulant protein domains -
 XX
 PS Disclosure; Fig 16; 197pp; English.
 XX
 CC The present sequence is the Ancylostoma ceylanicum nematode-extracted
 CC anticoagulant protein AcenAP4d2. Proteins of this kind have been shown to
 CC be effective at preventing blood clotting without causing excessive
 CC bleeding. The protein can be used in blood collection tubes to aid the
 CC isolation of plasma from the blood, to prevent thrombosis which may be
 CC linked to the rupture of an atherosclerotic plaque, acute myocardial
 CC infarction, angina, thrombolytic therapy, percutaneous transluminal
 CC coronary angioplasty, disseminated intravascular coagulopathy, infection,
 CC cancer and septic shock, and to produce antibodies. In the latter
 CC instance, the antibodies can be raised in order to detect infection by
 CC nematodes or as diagnostic tests. The protein can also be used as a
 CC vaccine against nematode parasites.
 XX
 SQ Sequence 82 AA;
 Query Match 93.8%; Score 30; DB 21; Length 82;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3 GFYRN 7
 Db 51 GFYRN 55
 |||||
 RESULT 13
 AAY30425
 ID AAY30425 standard; Protein; 83 AA.
 XX
 AC AAY30425;
 XX
 DT 15-NOV-1999 (first entry)
 XX
 DE Mature nematode extracted anticoagulant protein AduNAP7di.
 XX
 KW Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX
 OS Ancylostoma duodenale.
 XX
 PN US5955294-A.
 XX
 PD 21-SEP-1999.
 XX
 PF 19-APR-1996; 96US-0634641.
 XX
 PR 19-APR-1996; 96US-0634641.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.
 PR 05-JUN-1995; 95US-0486397.
 PR 05-JUN-1995; 95US-0486399.
 PR 17-OCT-1995; 95WO-US13231.
 XX
 PA (CORV-) CORVAS INT INC.
 XX
 PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
 PI Vlasuk GP;
 XX
 DR WPI; 1999-539569/45.
 XX
 PT Screening an isolated protein for Nematode-extracted Anticoagulant

PT Protein domains
XX
PS Disclosure; Columns 137-138; 197pp; English.
XX
CC The present sequence represents a nematode extracted anticoagulant protein (NAP). The protein has activity as an anticoagulant and/or serine protease inhibitor. The protein contains at least one NAP domain which has selective inhibitory activity for factor VIIa/TF. The specification describes a method for screening an isolated protein at least one domain for factor VIIa/TF selective inhibitory activity. The method comprises determining the time to clotting effected by a concentration of the isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo activated partial thromboplastin time (aPTT) assay; calculating prolongation of clotting effected by the isolated protein in each of the PT and aPTT assay, with respect to a baseline clotting value for each assay, where prolongation of clotting is calculated as fold elevation of clotting time relative to a baseline clotting value, where a doubling of clotting time is deemed a two-fold elevation; and calculating a PT to aPTT prolongation ratio, where a ratio at least one is indicative of factor VIIa/TF inhibitory activity. The method is useful for determining if a protein has factor VIIa/TF inhibitory activity.
XX
SQ Sequence 83 AA;
Query Match 93.8%; Score 30; DB 20; Length 83;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GFYRN 7
DB 55 GFYRN 59
|||||
RESULT 14
AAY30423
ID AAY30423 standard; Protein; 84 AA.
XX
AC AAY30423;
XX
DT 15-NOV-1999 (first entry)
XX
DE Mature nematode extracted anticoagulant protein AcanAP45d1.
XX
DE Nematode extracted anticoagulant protein; NAP; anticoagulant;
KW serine protease inhibitor; NAP domain; factor VIIa/TF.
XX
OS Ancylostoma caninum.
XX
PN US5955294-A.
XX
PD 21-SEP-1999.
XX
PF 19-APR-1996; 96US-0634641.
XX
PR 19-APR-1996; 96US-0634641.
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
PR 17-OCT-1995; 95WO-US13231.
XX
PA (CORV-) CORVAS INT INC.
XX
PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
PI Vlaeuk GP;
XX
DR WPI; 1999-539569/45.
XX
PT Screening an isolated protein for Nematode-extracted Anticoagulant Protein domains
XX

XX
PS Disclosure; Columns 135-136; 197pp; English.
XX
CC The present sequence represents a nematode extracted anticoagulant protein (NAP). The protein has activity as an anticoagulant and/or serine protease inhibitor. The protein contains at least one NAP domain which has selective inhibitory activity for factor VIIa/TF. The specification describes a method for screening an isolated protein at least one domain for factor VIIa/TF selective inhibitory activity. The method comprises determining the time to clotting effected by a concentration of the isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo activated partial thromboplastin time (aPTT) assay; calculating prolongation of clotting effected by the isolated protein in each of the PT and aPTT assay, with respect to a baseline clotting value for each assay, where prolongation of clotting is calculated as fold elevation of clotting time relative to a baseline clotting value, where a doubling of clotting time is deemed a two-fold elevation; and calculating a PT to aPTT prolongation ratio, where a ratio at least one is indicative of factor VIIa/TF inhibitory activity. The method is useful for determining if a protein has factor VIIa/TF inhibitory activity.
XX
SQ Sequence 84 AA;
Query Match 93.8%; Score 30; DB 20; Length 84;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GFYRN 7
DB 56 GFYRN 60
|||||
RESULT 15
AAY30424
ID AAY30424 standard; Protein; 84 AA.
XX
AC AAY30424;
XX
DT 15-NOV-1999 (first entry)
XX
DE Mature nematode extracted anticoagulant protein AcanAP47d1.
XX
DE Nematode extracted anticoagulant protein; NAP; anticoagulant;
KW serine protease inhibitor; NAP domain; factor VIIa/TF.
XX
OS Ancylostoma caninum.
XX
PN US5955294-A.
XX
PD 21-SEP-1999.
XX
PF 19-APR-1996; 96US-0634641.
XX
PR 19-APR-1996; 96US-0634641.
PR 18-OCT-1994; 94US-0326110.
PR 05-JUN-1995; 95US-0461965.
PR 05-JUN-1995; 95US-0465380.
PR 05-JUN-1995; 95US-0486397.
PR 05-JUN-1995; 95US-0486399.
PR 17-OCT-1995; 95WO-US13231.
XX
PA (CORV-) CORVAS INT INC.
XX
PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
PI Vlaeuk GP;
XX
DR WPI; 1999-539569/45.
XX
PT Screening an isolated protein for Nematode-extracted Anticoagulant Protein domains
XX

PS Disclosure; Columns 135-138; 197pp; English.

XX
CC The present sequence represents a nematode extracted anticoagulant
CC protein (NAP). The protein has activity as an anticoagulant and/or serine
CC protease inhibitor. The protein contains at least one NAP domain which
CC has selective inhibitory activity for factor VIIa/TF. The specification
CC describes a method for screening an isolated protein at least one domain
CC for factor VIIa/TF selective inhibitory activity. The method comprises
CC determining the time to clotting effected by a concentration of the
CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
CC activated partial thromboplastin time (aPTT) assay; calculating
CC prolongation of clotting effected by the isolated protein in each of
CC the PT and aPTT assay, with respect to a baseline clotting value for
CC each assay, where prolongation of clotting is calculated as fold
CC elevation of clotting time relative to a baseline clotting value, where
CC a doubling of clotting time is deemed a two-fold elevation; and
CC calculating a PT to aPTT prolongation ratio, where a ratio at least
CC one is indicative of factor VIIa/TF inhibitory activity. The method is
CC useful for determining if a protein has factor VIIa/TF inhibitory
CC activity.

XX
SQ Sequence 84 AA;

Query Match 93.8%; Score 30; DB 20; Length 84;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
|||
Db 56 GFYRN 60

Search completed: December 8, 2003, 09:16:39
Job time : 10.5455 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 1.83838 Seconds
(without alignments)
179.063 Million cell updates/sec

Title: US-09-498-556C-79
Perfect score: 32
Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	121	1 MTR3 MOUSE	Q92d0 mus musculus
2	30	93.8	175	1 HPRT_MYCGE	P47696 mycoplasma
3	30	93.8	175	1 HPRT_MYCPN	P75119 mycoplasma
4	30	93.8	245	1 END3 MYCTU	O69642 mycobacteri
5	30	93.8	279	1 UVEN MYCLU	P46303 micrococcus
6	30	93.8	361	1 SERC_XANAC	Q8p1y7 xanthomonas
7	30	93.8	361	1 SERC_XANCP	Q8p897 xanthomonas
8	30	93.8	361	1 SERC_YERPE	Q8zgb4 yersinia pe
9	30	93.8	362	1 SERC_XYLFA	Q9pb19 xylella fas
10	30	93.8	378	1 SERC_RALSO	Q8y0z0 ralstonia s
11	30	93.8	423	1 PUR2 RHIME	Q92r10 rhizobium m
12	30	93.8	481	1 SYW_MOUSE	P32921 mus musculu
13	30	93.8	533	1 TYRO_MOUSE	P11344 mus musculu
14	30	93.8	824	1 YOT5_CAEEL	P34651 caenorhabdi
15	30	93.8	900	1 XPC_MOUSE	P51612 mus musculu
16	30	93.8	939	1 XPC_HUMAN	Q01831 homo sapien
17	30	93.8	1564	1 PDRA_YEAST	P51533 saccharomyc
18	27	84.4	59	1 RL32_BORBU	O51646 borrelia bu
19	27	84.4	201	1 YAFJ_LACLA	Q9c1b0 lactococcus
20	27	84.4	218	1 KAD_CHLTE	Q8k669 chlorobium
21	27	84.4	234	1 DEOD_BUCAI	P57606 buchnera ap
22	27	84.4	299	1 YF35_METJA	Q58930 methanococc
23	27	84.4	318	1 SOLR_CIOAB	P33746 clostridium
24	27	84.4	320	1 K6PF_BUCAP	Q8k9n0 buchnera ap
25	27	84.4	340	1 ID17_ARCFU	Q28952 archaeoglob
26	27	84.4	344	1 Y613_METJA	Q58030 methanococc
27	27	84.4	465	1 YCJX_ECOLI	P76046 escherichia
28	27	84.4	471	1 SYW_HUMAN	P23381 homo sapien
29	27	84.4	471	1 YAL5_SCHPO	Q09928 schizosach
30	27	84.4	475	1 SYW_BOVIN	P17248 bos taurus
31	27	84.4	475	1 SYW_RABIT	P23612 cryptotagus
32	27	84.4	492	1 ANKH_MOUSE	Q9jh22 mus musculu
33	27	84.4	492	1 ANKH_RAT	P58366 rattus norv

34 27 84.4 550 1 PHNL_DESGI P12944 desulfovibr
35 27 84.4 608 1 SYT_PYRAE Q82wk4 pyrobaculum
36 27 84.4 667 1 TSI1_GIALA Q03185 giardia lam
37 27 84.4 713 1 TSM4_GIALA P21849 giardia lam
38 26 81.2 91 1 V179_FOWPV Q9J554 fowlpox vir
39 26 81.2 194 1 RK9_PEA P11894 pisum sativ
40 26 81.2 197 1 RK9_ARATH P25864 arabidopsis
41 26 81.2 218 1 YAFJ_LACLA Q9c1b5 lactococcus
42 26 81.2 219 1 END3_SYNV3 P71715 synchocyst
43 26 81.2 245 1 END3_MYCLE Q9cb92 mycobacteri
44 26 81.2 281 1 YQAC_LACLA Q9c1b9 lactococcus
45 26 81.2 282 1 YF57_STRPN Q97pp6 streptococc

ALIGNMENTS

RESULT 1
MTR3_MOUSE
ID MTR3_MOUSE STANDARD; PRT; 121 AA.
AC Q92d0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Myotubularin-related protein 3 (Fragment).
GN MTR3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98409499; PubMed=9736772;
RA Laporte J., Buj-Bello A., Tentler D., Kretz C., Dahl N.,
RA Mandel J.-L.;
RT "Characterization of the myotubularin dual specificity phosphatase
RT gene family from yeast to human."
RL Hum. Mol. Genet. 7:1703-1712(1998).
CC -!- FUNCTION: NOT KNOWN. COULD BE A TYROSINE-PHOSPHATASE.
CC -!- SIMILARITY: BELONGS TO THE MYOTUBULARIN-LIKE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF073881; AAC80003.1; -
CC MGD; MGI:1921552; Mtr3.
CC InterPro; IPR000387; TYR_phosphatase.
CC PROSITE; PS00383; TYR_PHOSPHATASE_1; PARTIAL.
CC PROSITE; PS00056; TYR_PHOSPHATASE_2; PARTIAL.
CC KW Hydrolase.
CC FT NON TER 1
SQ SEQUENCE 121 AA; 14292 MW; 0ECA07469DABB98 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 121;
Best Local Similarity 100.0%; Pred.No. 8;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 105 GFYRN 109

RESULT 2
HPRT_MYCGE
ID HPRT_MYCGE STANDARD; PRT; 175 AA.
AC P47696;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)

```
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8) (HGPRT)
DE (HGPRTase).
GN HPT OR MG458.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
RA Nguyen D.T., Dougherty T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Utherback T.A., Bost K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RL "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 9-98 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RA "A survey of the Mycoplasma genitalium genome by using random
RT sequencing.";
RL J. Bacteriol. 175:7918-7930(1993).
CC -|- CATALYTIC ACTIVITY: IMP + diphosphate = hypoxanthine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -|- CATALYTIC ACTIVITY: GMP + diphosphate = guanine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -|- PATHWAY: Purine salvage.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39727; AAC72478.1; -.
DR EMBL; U02193; AAD12479.1; -.
DR PIR; F64250; F64250.
DR HSP; P51900; LHGX.
DR TIGR; MG458; -.
DR InterPro; IPR005904; Hxn_phospho_trans.
DR InterPro; IPR002375; Pr/py_rp_transf.
DR InterPro; IPR000836; Prttransferase.
DR Pfam; PF00156; Pribosyltran; 1.
DR TIGRFAMs; TIGR01203; HGPRTase; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
KW Transferase; Glycosyltransferase; Purine salvage; Magnesium;
FT METAL 96 96 MAGNESIUM (BY SIMILARITY).
FT METAL 97 97 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 175 AA; 19759 MW; 9647332FF3E18E0 CRC64;
Query Match 93.8%; Score 30; DB 1; Length 175;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 GFYRN 7
Db 159 GFYRN 163
|||||
RESULT 3
HPRT_MYCPN STANDARD; PRT; 175 AA.
ID HPRT_MYCPN STANDARD; PRT; 175 AA.
```

```
AC P75119;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8) (HGPRT)
DE (HGPRTase).
GN HPT OR MPN672 OR MP170.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plegens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RL pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -|- CATALYTIC ACTIVITY: IMP + diphosphate = hypoxanthine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -|- CATALYTIC ACTIVITY: GMP + diphosphate = guanine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -|- PATHWAY: Purine salvage.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A8000018; AAB95818.1; -.
DR PIR; S73496; S73496.
DR HSP; P00492; IHMP.
DR InterPro; IPR005904; Hxn_phospho_trans.
DR InterPro; IPR002375; Pr/py_rp_transf.
DR InterPro; IPR000836; Prttransferase.
DR Pfam; PF00156; Pribosyltran; 1.
DR TIGRFAMs; TIGR01203; HGPRTase; 1.
DR PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
KW Transferase; Glycosyltransferase; Purine salvage; Magnesium;
KW Complete proteome.
FT METAL 96 96 MAGNESIUM (BY SIMILARITY).
FT METAL 97 97 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 175 AA; 19620 MW; 1E5A9FAA01D69854 CRC64;
Query Match 93.8%; Score 30; DB 1; Length 175;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 GFYRN 7
Db 159 GFYRN 163
|||||
RESULT 4
END3_MYCTU STANDARD; PRT; 245 AA.
ID END3_MYCTU STANDARD; PRT; 245 AA.
AC O69642;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Endonuclease III (EC 4.2.99.18) (DNA-(apurinic or apyrimidinic site)
DE lyase).
DE lyase).
GN NTH OR RV3674C OR MT3775 OR MTV025.022c.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
```

OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Unayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Blehni W.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: HAS BOTH AN APURINIC AND/OR APYRIMIDINIC ENDONUCLEASE
 CC ACTIVITY AND A DNA N-GLYCOSYLASE ACTIVITY. INCISES DAMAGED STRAND, 5'
 CC CYTOSINES, THYMINES AND GUANINES. ACTS ON A DAMAGED STRAND, 5'
 CC FROM THE DAMAGED SITE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: The C-O-P bond 3' to the apurinic or
 CC apyrimidinic site in DNA is broken by a beta-elimination reaction,
 CC leaving a 3'-terminal unsaturated sugar and a product with a
 CC terminal 5'-phosphate.
 CC -1- COFACTOR: BINDS A 4FE-4S CLUSTER WHICH IS NOT IMPORTANT FOR THE
 CC CATALYTIC ACTIVITY, BUT WHICH IS PROBABLY INVOLVED IN THE PROPER
 CC POSITIONING OF THE ENZYME ALONG THE DNA STRAND (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NTH/MUTY FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL022121; CRA17996.1; ALT INIT.
 CC EMBL; AE007175; AAK48142.1; ALT_INIT.
 CC HSSP; P20625; 2ABK.
 CC TIGR; MT3775; -;
 CC Tuberculaist; RV3674c; -;
 CC InterPro; IPR003265; Endo_3c.
 CC InterPro; IPR004035; EndoIII_FCL.
 CC InterPro; IPR004036; EndoIII_Hhh.
 CC InterPro; IPR003651; Fes_bind.
 CC InterPro; IPR000445; Hhh.
 CC InterPro; IPR003583; Hhh_1.
 CC InterPro; IPR005759; Nth.
 CC Pfam; PF00730; Hhh-GPD; 1.
 CC Pfam; PF00633; Hhh; 1.
 CC SMART; SM00478; ENDO3c; 1.
 CC SMART; SM00525; FES; 1.
 CC SMART; SM00278; Hhh1; 1.
 CC TIGRFAMs; TIGR01083; nth; 1.
 CC PROSITE; PS00764; ENDONUCLEASE_III_1; 1.
 CC PROSITE; PS01155; ENDONUCLEASE_III_2; 1.
 CC Hydrolase; Nuclease; Endonuclease; Multifunctional enzyme; DNA repair;
 KW Glycosidase; Lyase; Iron-sulfur; 4Fe-4S; Complete proteome.
 FT METAL 198 198 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 205 205 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 208 208 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

FT METAL 214 214 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 SQ SEQUENCE 245 AA; 27030 MW; 2B6D16195DD090DE CRC64;
 Query Match 93.8%; Score 30; DB 1; Length 245;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 GFYRN 7
 Db |||||
 91 GFYRN 95
 RESULT 5
 ID JVEN MICLU STANDARD; PRT; 279 AA.
 AC P46303;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ultraviolet N-glycosylase/AP lyase (UV-endonuclease) (Pyrimidine dimer
 DE glycosylase).
 GN PDG.
 OS Micrococcus luteus (Micrococcus lysodeikticus).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococciaceae; Micrococaceae; Micrococcus.
 RN [1]_TaxID=1270;
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-35.
 RC STRAIN=ATCC 4698;
 RX MEDLINE=96007490; PubMed=7559510;
 RA Pierson C.E., Prince M.A., Augustine M.L., Dodson M.L., Lloyd R.S.;
 RT "Purification and cloning of Micrococcus luteus ultraviolet
 RT endonuclease, an N-glycosylase/abasic lyase that proceeds via an
 RT imino enzyme-DNA intermediate.";
 RL J. Biol. Chem. 270:23475-23484(1995).
 CC -1- FUNCTION: HAS BOTH, AN APURINIC AND/OR APYRIMIDINIC ENDONUCLEASE
 CC ACTIVITY AND A DNA N-GLYCOSYLASE ACTIVITY. INITIATES REPAIR AT
 CC CIS-SYN PYRIMIDINE DIMERS. PROCEEDS VIA AN IMINO ENZYME:DNA
 CC INTERMEDIATE.
 CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR UAG OCCURS BETWEEN
 CC CODONS FOR GLY-268 AND ALA-270. TWO FORMS OF 31 KDA AND 32 KDA
 CC HAVE BEEN DETECTED.
 CC -1- SIMILARITY: BELONGS TO THE NTH/MUTY FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U22181; AAA86508.1; ALT_TERM.
 CC HSSP; P20625; 2ABK.
 CC InterPro; IPR003265; Endo_3c.
 CC InterPro; IPR004035; EndoIII_FCL.
 CC InterPro; IPR004036; EndoIII_Hhh.
 CC InterPro; IPR003651; Fes_bind.
 CC InterPro; IPR000445; Hhh.
 CC InterPro; IPR003583; Hhh_1.
 CC InterPro; IPR005759; Nth.
 CC Pfam; PF00730; Hhh-GPD; 1.
 CC Pfam; PF00633; Hhh; 1.
 CC SMART; SM00478; ENDO3c; 1.
 CC SMART; SM00525; FES; 1.
 CC SMART; SM00278; Hhh1; 1.
 CC TIGRFAMs; TIGR01083; nth; 1.
 CC PROSITE; PS00764; ENDONUCLEASE_III_1; 1.
 CC PROSITE; PS01155; ENDONUCLEASE_III_2; 1.
 CC Hydrolase; Nuclease; Endonuclease; DNA repair; Glycosidase;
 KW Hydrolyase; Nuclease; Endonuclease; DNA repair; Glycosidase;
 FT CHAIN 1 268 UV ENDONUCLEASE 31 kDa FORM.
 FT CHAIN 1 279 UV ENDONUCLEASE 32 kDa FORM (PROBABLE).
 FT CHAIN

```
FT METAL 203 203 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 210 210 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 213 213 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 219 219 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 279 AA; 30469 MW; C385369A1827C005 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 279;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 95 GFYRN 99

RESULT 6
SERC_XANAC STANDARD; PRT; 361 AA.
AC Q8PLV7;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).
GN SERC OR XAC1648.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -|- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
CC phosphonooxypyruvate + L-glutamate.
CC -|- COPACTOR: Pyridoxal phosphate (By similarity).
CC -|- PATHWAY: Required both in major phosphorylated pathway of serine
CC biosynthesis and in the biosynthesis of pyridoxine (By
CC similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to class-V of pyridoxal-phosphate-dependent
CC aminotransferases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE011797; AAM36516.1; -.
CC HAMAP; MF_00160; -.
CC InterPro; IPR000192; Aminotransf.
CC InterPro; IPR003248; Pser aminotransf.
CC Pfam; PF00266; aminotran_5; 1.
```

```
DR ProDom; PD001544; Pser aminotransf; 1.
DR TIGRFAMs; TIGR01364; serC1; 1.
DR PROSITE; PS00595; AA_TRANSFER_CLASS_5; 1.
KW Serine biosynthesis; Pyridoxine biosynthesis; Transferase;
KW Aminotransferase; Pyridoxal phosphate; Complete proteome.
FT BINDING 196 196 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 361 AA; 38680 MW; 75EC2B319C493982 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7
Db 283 GFYRN 287

RESULT 7
SERC_XANCP STANDARD; PRT; 361 AA.
AC Q8PA97;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).
GN SERC OR XCC1589.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
CC -|- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
CC phosphonooxypyruvate + L-glutamate.
CC -|- COPACTOR: Pyridoxal phosphate (By similarity).
CC -|- PATHWAY: Required both in major phosphorylated pathway of serine
CC biosynthesis and in the biosynthesis of pyridoxine (By
CC similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: Belongs to class-V of pyridoxal-phosphate-dependent
CC aminotransferases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE012260; AAM40884.1; -.
CC HAMAP; MF_00160; -.
CC InterPro; IPR000192; Aminotransf.
CC InterPro; IPR003248; Pser aminotransf.
```


DR InterPro; IPR003248; Peer aminotransf.
DR Pfam; PF00266; aminotran 5; 1.
DR ProDom; PD001544; Peer aminotransf; 1.
DR TIGRFAMs; TIGR01364; serC 1; 1.
DR PROSITE; PS00595; AA TRANSFER CLASS 5; 1.
KW Serine biosynthesis; Pyridoxine biosynthesis; Transferase;
KW Aminotransferase; Pyridoxal phosphate; Complete proteome.
FT BINDING 196 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 361 AA; 38759 MW; F0025FFB9BF7B65 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 283 GFYRN 287
|||||

RESULT 8
SERC_YERPE STANDARD; PRT; 361 AA.
AC Q8ZG84;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).
GN SERC OR YP01389 OR Y2784.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Lies P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Milles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM";
RL J. Bacteriol. 184:4601-4611(2002).
CC -!- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
phosphonooxypyruvate + L-glutamate.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: Required both in major phosphorylated pathway of serine
biosynthesis and in the biosynthesis of pyridoxine (By
similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to class-V of pyridoxal-phosphate-dependent
aminotransferases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AJ414148; CAC90218.1; -;
DR EMBL; AE013881; AM86336.1; -;
DR FIR; AG0169; AG0169.
DR HAMAP; MF_00160; -; 1.
DR InterPro; IPR00192; Aminotransfv.
DR InterPro; IPR003248; Peer aminotransf.
DR Pfam; PF00266; aminotran 5; 1.
DR ProDom; PD001544; Peer aminotransf; 1.
DR TIGRFAMs; TIGR01364; serC 1; 1.
DR PROSITE; PS00595; AA TRANSFER CLASS 5; 1.
KW Serine biosynthesis; Pyridoxine biosynthesis; Transferase;
KW Aminotransferase; Pyridoxal phosphate; Complete proteome.
FT BINDING 197 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 361 AA; 40083 MW; 8827EF1419782D88 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 283 GFYRN 287
|||||

RESULT 9
SERC_XYLFA STANDARD; PRT; 362 AA.
AC Q9PB19;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Phosphoserine aminotransferase (EC 2.6.1.52) (PSAT).
GN SERC OR XF2326.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5c;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Aruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Colauto N.B., Cristofani M., Dias-Neto E., Docena C., El-Dorfi H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA de Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159(2000).
CC -!- CATALYTIC ACTIVITY: O-phospho-L-serine + 2-oxoglutarate = 3-
phosphonooxypyruvate + L-glutamate.
CC -!- COFACTOR: Pyridoxal phosphate (By similarity).

```

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; AL646061; CAD14605.1; -.
HAMAP; MF_00160; -. 1.
InterPro; IPR000192; Aminotransfv.
InterPro; IPR003248; Pser_aminotransf.
Pfam; PF00266; aminotran_5; 1.
ProDom; PD001544; Pser_aminotransf; 1.
TIGRFAMs; TIGR01364; serC_1; 1.
PROSITE; PS00595; AA_TRANSF_CLASS_5; FALSE_NEG.
Serine biosynthesis; Pyridoxine biosynthesis; transferase;
Aminotransferase; Pyridoxal phosphate; Complete proteome.
BINDING 214 214 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SEQUENCE 378 AA; 41860 MW; D4CA356D9098E97B CRC64;
-----
Query Match 93.8%; Score 30; DB 1; Length 378;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
-----
QY 3 GFYRN 7
DB 300 GFYRN 304
|||||
-----
RESULT 11
PUR2_RHIME STANDARD; PRT; 423 AA.
ID PUR2_RHIME
AC Q92RJO;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoribosylamine-glycine ligase (EC 6.3.4.13) (GARS) (Glycinamide
DE ribonucleotide synthetase) (Phosphoribosylglycinamide synthetase).
DE PURD OR R00858 OR SMC00993.
DE Rhizobium melloti (Sinorhizobium melloti).
DE Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
DE Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
DE NCBI_TaxID=382;
DE [1]_TaxID=382;
DE SEQUENCE FROM N.A.
DE STRAIN=2013;
DE MEDLINE=121396507; Pubmed=11481430;
DE Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
DE Boisdard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
DE Godrie T., Goffeau A., Kahn D., Kiss E., Leilaure V., Masuy D.,
DE Pohl T., Portetelle D., Puchler A., Purnelle B., Ramsperger U.,
DE Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
DE "Analysis of the chromosome sequence of the legume symbiont
DE Sinorhizobium melloti strain 1021."
DE Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DE -!- CATALYTIC ACTIVITY: ATP + 5-phospho-D-ribosylamine + glycine = ADP
DE + phosphate + N(1)-(5-phospho-D-ribosyl)glycinamide.
DE -!- PATHWAY: De novo purine biosynthesis; second step.
DE -!- SIMILARITY: Belongs to the GARS family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; AL591785; CAC45430.1; -.
HAMAP; MF_00138; -. 1.
InterPro; IPR000115; Gars.

```

```

DR Pfam; PF01071; GARS; 1.
DR DR PF02842; GARS B; 1.
DR DR PF02843; GARS C; 1.
DR DR PF02844; GARS N; 1.
DR TIGRFAMs; TIGR00877; purD; 1.
DR PROSITE; PS00184; GARS; 1.
KW Purine biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 423 AA; 44324 MW; 5E65E13B606D204B CRC64;

Query Match 93.8%; Score 30; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 407 GFYRN 411

RESULT 12
SYW MOUSE
ID SYW_MOUSE STANDARD; PRT; 481 AA.
AC P32921;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Typtophanyl-tRNA synthetase (EC 6.1.1.2) (Typtophan--trna ligase)
DE (TrpRS).
DE WARS OR WRS.
GN MUS MUSCULUS (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=95018226; PubMed=7932716;
RA Fajot B., Sarger C., Bonnet J., Garret M.;
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-
RT tRNA synthetase in murine embryonic stem cells.";
RL J. Mol. Biol. 242:599-603(1994).
CC -!- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=Long;
CC IsoID=P32921-1; Sequence=Displayed;
CC Name=2; Synonyms=Short;
CC IsoID=P32921-2; Sequence=VSP_006313;
CC -!- TISSUE SPECIFICITY: Isoform 2 is widely expressed, isoform 1 is
CC found only in embryonic stem cells.
CC -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC -!- SIMILARITY: Contains 1 WHEP-TRS domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69656; CAA49347.1; -.
DR EMBL; X69657; CAA49348.1; -.
DR PIR; S50053; S50053.
DR MGI; MGI:104630; Wars.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; Trp tRNA-synt_1b.
DR InterPro; IPR000738; WHEP-TRS.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR Pfam; PF00458; WHEP-TRS; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR TIGRFAMs; TIGR00233; trpS; 1.

DR PROSITE; PS00178; AA TRNA LIGASE_I; 1.
DR PROSITE; PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSKS" REGION.
FT VARSPLIC 476 481 Missing (in isoform 2).
FT FTID=VSP_006313.
SQ SEQUENCE 481 AA; 54282 MW; B05A452C08074F52 CRC64;

Query Match 93.8%; Score 30; DB 1; Length 481;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 250 GFYRN 254

RESULT 13
TYRO MOUSE
ID TYRO_MOUSE STANDARD; PRT; 533 AA.
AC P11344;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tyrosinase precursor (EC 1.14.18.1) (Monophenol monooxygenase)
DE (Albino locus protein).
GN TYR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2J;
RX MEDLINE=88268910; PubMed=3134020;
RA Kwon B.S., Wakulchik M., Haq A.K., Halaban R., Kestler D.;
RT "Sequence analysis of mouse tyrosinase cDNA and the effect of
RT melanotropin on its gene expression.";
RL Biochem. Biophys. Res. Commun. 153:1301-1309(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Himalayan;
RX MEDLINE=89273644; PubMed=2567165;
RA Kwon B.S., Halaban R., Chintamani C.;
RT "Molecular basis of mouse Himalayan mutation.";
RL Biochem. Biophys. Res. Commun. 161:252-260(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89030636; PubMed=3141148;
RA Mueller G., Ruppert S., Schmid E., Schuetz G.;
RT "Functional analysis of alternatively spliced tyrosinase gene
RT transcripts.";
RL EMBO J. 7:2723-2730(1988).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=89193679; PubMed=2494997;
RA Terao M., Tabé L., Garattini E., Sartori D., Studer M., Mintz B.;
RT "Isolation and characterization of variant cDNAs encoding mouse
RT tyrosinase.";
RL Biochem. Biophys. Res. Commun. 159:848-853(1989).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Yamamoto H., Takeuchi S., Kudo T., Makino K., Nakata A., Shinoda T.,
RA Takeuchi T.;
RT "Cloning and sequencing of mouse tyrosinase cDNA.";
RL Jpn. J. Genet. 62:271-274(1987).
RN [6]
RP SEQUENCE OF 1-273 FROM N.A.
RX MEDLINE=90212084; PubMed=2517217;

```

RA Yamamoto H., Takeuchi S., Kudo T., Sato C., Takeuchi T.;
RT "Melanin production in cultured albino melanocytes transfected with
RL mouse tyrosinase cDNA."; Jpn. J. Genet. 64:121-135 (1989).
RN [7]
RP VARIANT ALBINO.
RC STRAIN=BALB/c;
RX MEDLINE=90249393; PubMed=2110899;
RA Shibahara S., Okinaga S., Tomita Y., Takeda A., Yamamoto H., Sato M.,
RA Takeuchi T.;
RT "A point mutation in the tyrosinase gene of BALB/c albino mouse
causing the cysteine--serine substitution at position 85."; Eur. J. Biochem. 189:455-461 (1990).
RN [8]
RP VARIANT CHINCHILLA MICE.
RX MEDLINE=90360993; PubMed=2118105;
RA Beermann F., Ruppert S., Hummler E., Bosch F.X., Mueller G.,
RA Ruether U., Schuetz G.;
RT "Rescue of the albino phenotype by introduction of a functional
tyrosinase gene into mice."; EMBO J. 9:2819-2826 (1990).
RL
CC -!- FUNCTION: THIS IS A COPPER-CONTAINING OXIDASE THAT FUNCTIONS IN
THE FORMATION OF PIGMENTS SUCH AS MELANINS AND OTHER POLYPHENOLIC
COMPOUNDS. CATALYZES THE RATE-LIMITING CONVERSIONS OF TYROSINE TO
DOPA. DOPA TO DOPA-QUINONE AND POSSIBLY 5,6-DIHYDROXYINDOLE TO
INDOLE-5,6 QUINONE.
CC -!- CATALYTIC ACTIVITY: L-tyrosine + L-DOPA + O(2) = L-DOPA +
DOPAquinone + H(2)O.
CC -!- COFACTOR: Binds 2 copper ions per subunit.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Melanosomal.
CC -!- DISEASE: DEFECTS IN TYR RESULT IN VARIOUS FORMS OF ALBINISM.
CC HIMALAYAN STRAIN TYROSINASE IS TEMPERATURE-SENSITIVE.
CC -!- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
CC -!- CAUTION: REF.4 SEQUENCE WAS INCORRECT DUE TO A DELETION OF EXON 3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; D00440; BAA00341.1; -;
DR EMBL; M20234; AAA40516.1; -;
DR EMBL; M26729; AAA37806.1; -;
DR EMBL; X12782; CAA31273.1; -;
DR EMBL; M24560; AAA40517.1; -;
DR EMBL; D00131; BAA00079.1; -;
DR EMBL; X51743; CAA36033.1; -;
DR EMBL; D00439; BAA00340.1; -;
DR PIR; A27711; YRMSCS.
DR MGD; MGI:98880; Tyr.
DR InterPro; IPR002227; Tyrosinase.
DR Pfam; PF00264; Tyrosinase; 1.
DR PRINTS; PF00092; TYROSINASE.
DR PROSITE; PS00497; TYROSINASE_1; 1.
DR PROSITE; PS00498; TYROSINASE_2; 1.
KW Oxidoreductase; Monooxygenase; Copper; Glycoprotein; Signal;
KW Transmembrane; Melanin biosynthesis; Disease mutation; Albinism.
FT SIGNAL 1 18
FT CHAIN 19 533 TYROSINASE.
FT DOMAIN 19 476 LUMENAL, MELANOSOME (POTENTIAL).
FT TRANSMEM 477 497 POTENTIAL.
FT DOMAIN 498 533 CYTOPLASMIC (POTENTIAL).
FT METAL 180 180 COPPER A (BY SIMILARITY).
FT METAL 202 202 COPPER A (BY SIMILARITY).
FT METAL 211 211 COPPER A (BY SIMILARITY).
FT METAL 363 363 COPPER B (BY SIMILARITY).
FT METAL 367 367 COPPER B (BY SIMILARITY).
FT METAL 390 390 COPPER B (BY SIMILARITY).
FT METAL 503 508 POLY-LYS.
FT DOMAIN 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD

FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 103 103 C -> S (IN ALBINO MICE).
FT VARIANT 420 420 H -> R (IN HIMALAYAN STRAIN).
FT VARIANT 482 482 A -> T (IN CHINCHILLA MICE).
FT CONFLICT 40 40 M -> I (IN REF. 2).
FT CONFLICT 197 197 D -> Q (IN REF. 2).
FT CONFLICT 264 264 S -> I (IN REF. 3).
FT CONFLICT 346 346 V -> G (IN REF. 2, 3 AND 4).
FT CONFLICT 348 356 ASPLTGAD -> LPEHNGCEG (IN REF. 5).
FT CONFLICT 357 403 MISSING (IN REF. 5).
FT CONFLICT 471 495 ASRWFLWGLGALVGVIAAALSGL -> GQSYLAWSWGS
FT CONFLICT 496 533 TGSCYCCSSSLWA (IN REF. 5).
FT CONFLICT 533 533 MISSING (IN REF. 5).
SQ SEQUENCE 533 AA; 60648 MW; 4B711312DDB6F7D1 CRC64;
Query Match 93.8%; Score 30; DB 1; Length 533;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GFYRN 7
DB 458 GFYRN 462
RESULT 14
YOT5 CABEL STANDARD; PRT; 824 AA.
ID YOT5 CABEL
AC P34651;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ZK632.5 in chromosome III.
GN ZK632.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans."; Nature 368:32-38 (1994).
RL Nature 368:32-38 (1994).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; Z22181; CAA80182.1; -;
DR PIR; S40937; S40937.
DR WormPep; ZK632.5; CE00422.
KW Hypothetical protein.
SQ SEQUENCE 824 AA; 95726 MW; 05074058F5B73919 CRC64;

```
Query Match          93.8%; Score 30; DB 1; Length 824;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GFYRN 7
DB      206 GFYRN 210

RESULT 15
XPC_MOUSE STANDARD; PRT; 900 AA.
AC P51612; P54732;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE DNA-repair protein complementing XP-C cells homolog (Xeroderma
GN pigmentosum group C complementing protein homolog) (P125).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96184849; PubMed=8604333;
RA Li L., Peterson C., Legerski R.;
RT "Sequence of the mouse XPC cDNA and genomic structure of the human
RL XPC gene.";
RL Nucleic Acids Res. 24:1026-1028 (1996).
RN [2]
RP SEQUENCE OF 28-587 FROM N.A.
RC STRAIN=129/Sv;
RX MEDLINE=95405469; PubMed=7675084;
RA Sands A.T., Abuin A., Sanchez A., Conti C.J., Bradley A.;
RT "High susceptibility to ultraviolet-induced carcinogenesis in mice
RL lacking XPC.";
RL Nature 377:162-165 (1995).
CC -!- FUNCTION: INVOLVED IN DNA EXCISION REPAIR. MAY PLAY A PART IN DNA
CC DAMAGE RECOGNITION AND/OR IN ALTERING CHROMATIN STRUCTURE TO
CC ALLOW ACCESS BY DAMAGE-PROCESSING ENZYMES.
CC -!- SUBUNIT: HETERODIMER OF A 125 Kda SUBUNIT (P125) AND OF A
CC 58 Kda SUBUNIT (P58).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: SOME, TO YEAST RAD4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U27398; AAC52500.1; --
DR EMBL; U40005; AAA82720.1; --
DR PIR; S70630; S70630.
DR MGD; MGI:103557; Xpc.
DR GO; GO:0006289; P:nucleotide-excision repair; IMP.
DR InterPro; IPR004583; Rad4.
DR Pfam; PF03835; Rad4; 1.
DR TIGRfams; TIGR00605; rad4; 1.
KW DNA repair; DNA-binding; Nuclear protein.
FT CONFLICT 28 30 AVN -> CSD (IN REF. 2).
FT CONFLICT 53 53 S -> L (IN REF. 2).
FT CONFLICT 67 67 L -> F (IN REF. 2).
FT CONFLICT 70 70 L -> S (IN REF. 2).
FT CONFLICT 134 135 RG -> TP (IN REF. 2).
FT CONFLICT 165 170 EQVENM -> GVHEDT (IN REF. 2).
FT CONFLICT 181 181 S -> N (IN REF. 2).
FT CONFLICT 187 187 S -> N (IN REF. 2).
FT CONFLICT 190 190 R -> S (IN REF. 2).

Query Match          93.8%; Score 30; DB 1; Length 900;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 GFYRN 7
DB      182 GFYRN 186

Search completed: December 8, 2003, 09:14:13
Job time : 2.83838 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 7.35354 Seconds
(without alignments)
245.646 Million cell updates/sec

Title: US-09-498-556C-79
Perfect score: 32
Sequence: 1 XXGFYRN 7

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	77	Q8X3Q1	Q8x3q1 escherichia
2	30	93.8	88	Q8FIE2	Q8fie2 escherichia
3	30	93.8	91	Q16938	Q16938 ancylostoma
4	30	93.8	102	Q962V8	Q962v8 ancylostoma
5	30	93.8	110	Q9KYV8	Q9kyv8 streptomyce
6	30	93.8	134	P74345	P74345 synchocyst
7	30	93.8	154	Q8ITP8	Q8itp8 oesophagost
8	30	93.8	158	Q44490	Q44490 anabaena va
9	30	93.8	172	Q9A2A6	Q9a2a6 caulobacter
10	30	93.8	178	Q9TAJ9	Q9taj9 cafeteria r
11	30	93.8	184	Q9VGI2	Q9vgi2 drosophila
12	30	93.8	184	Q96723	Q96723 drosophila
13	30	93.8	191	Q8L803	Q8l803 triticum ae
14	30	93.8	216	Q8TNS2	Q8tns2 methanosarc
15	30	93.8	222	Q8DIE9	Q8die9 synchococc
16	30	93.8	223	Q8YQ65	Q8yq65 anabaena sp

17	30	93.8	228	Q8GMR5	Q8gmr5 synchococc
18	30	93.8	230	Q8UBS5	Q8ubz5 grapevine f
19	30	93.8	237	Q8HPV4	Q8hvp4 halobacteri
20	30	93.8	260	Q8NTL4	Q8ntl4 corynebacte
21	30	93.8	264	Q8FSU2	Q8fsu2 corynebacte
22	30	93.8	266	Q8ZTI9	Q8zt19 pyrobaculum
23	30	93.8	302	Q8C2V6	Q8c2v6 mus musculu
24	30	93.8	325	Q9BKX3	Q9bkx3 lucilia cup
25	30	93.8	354	Q9VEY0	Q9vey0 drosophila
26	30	93.8	361	Q8PLY7	Q8ply7 xanthomonas
27	30	93.8	361	Q8PA37	Q8pa37 xanthomonas
28	30	93.8	361	Q8ZGB4	Q8zgb4 yersinia pe
29	30	93.8	362	Q9PBI9	Q9pb19 xylella fas
30	30	93.8	378	Q8Y0Z0	Q8y0z0 ralstonia s
31	30	93.8	408	Q01482	Q01482 caenorhabdi
32	30	93.8	415	Q9U153	Q9u153 leishmania
33	30	93.8	465	Q9C8J9	Q9c8j9 arabidopsis
34	30	93.8	470	Q8EDG2	Q8edg2 shewanella
35	30	93.8	472	Q02868	Q02868 saccharomyc
36	30	93.8	475	Q9P6K0	Q9p6k0 schizosacch
37	30	93.8	475	Q9DC65	Q9dc65 mus musculu
38	30	93.8	481	Q99J58	Q99j58 mus musculu
39	30	93.8	511	Q8GRX1	Q8grx1 arabidopsis
40	30	93.8	533	Q91XK0	Q91xk0 mus musculu
41	30	93.8	554	Q8ZZL2	Q8zzl2 pyrobaculum
42	30	93.8	596	Q86468	Q86468 rhodobacter
43	30	93.8	600	Q9NL27	Q9nl27 ciona intes
44	30	93.8	611	Q9L775	Q9l775 arabidopsis
45	30	93.8	611	Q8GZT8	Q8gzt8 arabidopsis

ALIGNMENTS

RESULT 1

Q8X3Q1 PRELIMINARY; PRT; 77 AA.
AC Q8X3Q1;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DE 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein z1836.
GN z1836.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocous K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RL "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RT Nature 409:529-533(2001).
DR EMBL; AE005327; AAG55934.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 77 AA; 8087 MW; F85D56712A58545E CRC64;

Query Match 93.8%; Score 30; DB 16; Length 77;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 15 GFYRN 19

RESULT 2

```
Q8FIE2
ID Q8FIE2 PRELIMINARY; PRT; 88 AA.
AC Q8FIE2
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN C1484
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
DR EMBL; AE016759; AAN79953.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 88 AA, 9179 MW; 6709ABA0C8259B8B CRC64;

Query Match 93.8%; Score 30; DB 16; Length 88;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 15 GFYRN 19

RESULT 3
Q16938 PRELIMINARY; PRT; 91 AA.
ID Q16938
AC Q16938
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anti-coagulant protein C2 precursor (Fragment).
OS Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatoides; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98298519; PubMed=9634780;
RA Jespers L.S., Messens J.H., De Keyser A., Beckhout D.,
RA van den Brande I., Gansemans Y.G., Lauwereys M.J., Vlasuk G.P.,
RA Stanssens P.E.;
RT "Surface expression and ligand-based selection of cDNAs fused to
RT filamentous phage gene VI.";
RL Biotechnology 13:378-382(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312555; PubMed=8700900;
RA Stanssens P., Bergum P.W., Gansemans Y., Jaspers L., Laroche Y.,
RA Huang S., Maki S., Messens J., Lauwereys M., Cappello M., Hotez P.J.,
RA Lasters I., Vlasuk G.P.;
RT "Anticoagulant repertoire of the hookworm Ancylostoma caninum.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:2149-2154(1996).
DR EMBL; U30793; AAC47080.1; -.
DR HSSP; P56682; 1CCV.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002519; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
DR PROSITE; PS01186; EGF_2; 1.
KW SIGNAL.
FT SIGNAL.
FT CHAIN 1 19 POTENTIAL.
SQ SEQUENCE 102 AA; 11834 MW; 315722980EF723E7 CRC64;

Query Match 93.8%; Score 30; DB 5; Length 102;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 71 GFYRN 75

RESULT 5
Q9KYV8 PRELIMINARY; PRT; 110 AA.
ID Q9KYV8
AC Q9KYV8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SCO3187.
GN SCO3187 OR SCE22.04.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
```

```
FT SIGNAL <1 7 POTENTIAL.
FT CHAIN 8 91 ANTI-COAGULANT PROTEIN C2.
SQ SEQUENCE 91 AA; 10358 MW; ECBI1CB4597C24DA CRC64;

Query Match 93.8%; Score 30; DB 5; Length 91;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 62 GFYRN 66

RESULT 4
Q962V8 PRELIMINARY; PRT; 102 AA.
ID Q962V8
AC Q962V8
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Anticoagulant peptide-1 precursor.
GN ACEAP-1.
OS Ancylostoma ceylanicum.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatoides; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=53326;
RN [1]
RP SEQUENCE FROM N.A.
RA Harrison L.M., Cappello M.;
RT "Cloning of the major factor Xa inhibitor (AceAP-1) from Ancylostoma
RT ceylanicum";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF399710; AAK81733.1; -.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR002919; TIL_Cysrich.
DR Pfam; PF01826; TIL; 1.
DR PROSITE; PS01186; EGF_2; 1.
KW SIGNAL.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 102 ANTI-COAGULANT PEPTIDE-1.
SQ SEQUENCE 102 AA; 11834 MW; 315722980EF723E7 CRC64;

Query Match 93.8%; Score 30; DB 5; Length 102;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 71 GFYRN 75

RESULT 5
Q9KYV8 PRELIMINARY; PRT; 110 AA.
ID Q9KYV8
AC Q9KYV8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SCO3187.
GN SCO3187 OR SCE22.04.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
```


RA Rabbinitzsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RT Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)." (2002).
RT Nature 417:141-147 (2002).
DR EMBL; AL939115; CAB90971.1; -;
SQ SEQUENCE 110 AA; 10831 MW; 0893F0F52B8EA8FA CRC64;

Query Match 93.8%; Score 30; DB 16; Length 110;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 51 GFYRN 55

RESULT 6
P74345 PRELIMINARY; PRT; 134 AA.
ID P74345
AC P74345
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein slr1628.
GN SLR1628.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136 (1996).
DR EMBL; D90914; BAA18439.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 134 AA; 14545 MW; 7E5414E80FB43D10 CRC64;

Query Match 93.8%; Score 30; DB 16; Length 134;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 27 GFYRN 31

RESULT 7
Q81TP8 PRELIMINARY; PRT; 154 AA.
ID Q81TP8
AC Q81TP8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative trypsin-like inhibitor protein precursor.
GN MCRP.
OS Oesophagostomum dentatum.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Strongyloidea; Chabertiidae; Oesophagostomum.
OX NCBI_TaxID=61180;
RN [1]
RP SEQUENCE FROM N.A.
RA Boag P.R., Ranganathan S., Newton S.E., Gasser R.B.;

RT "Identification of a male-specific nematode protein with two trypsin
RT like inhibitor domains.";
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF399936; AAN32637.1; -;
KW Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 154
SQ SEQUENCE 154 AA; 16564 MW; A7F566E3957DA819 CRC64;

Query Match 93.8%; Score 30; DB 5; Length 154;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 64 GFYRN 68

RESULT 8
Q44490 PRELIMINARY; PRT; 158 AA.
ID Q44490
AC Q44490
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Orf2.
DE Anabaena variabilis.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=1172;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29413;
RX MEDLINE=96016168; PubMed=7568132;
RA Thiel T., Lyons E.M., Erker J.C., Ernst A.;
RT "A second nitrogenase in vegetative cells of a heterocyst-forming
RT cyanobacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:9358-9362 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29413;
RA Thiel T., Lyons E.M., Erker J.C.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U49859; AAA93026.1; -;
DR InterPro; IPR004952; DUF269.
DR Pfam; PF03270; DUF269; 1
SQ SEQUENCE 158 AA; 17791 MW; D157EBF59C36FEED CRC64;

Query Match 93.8%; Score 30; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 35 GFYRN 39

RESULT 9
Q9A2A6 PRELIMINARY; PRT; 172 AA.
ID Q9A2A6
AC Q9A2A6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein CC3660.
GN CC3660.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;

```

RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Winn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,
RA "Complete genome sequence of Caulobacter crescentus."
RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AEO06024; AAK25622.1; -.
DR TIGR; CC3660; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 172 AA; 18425 MW; 5DD52E712F406F6A CRC64;

Query Match 93.8%; Score 30; DB 16; Length 172;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 73 GFYRN 77

RESULT 10
Q9TAJ9 PRELIMINARY; PRT; 178 AA.
AC Q9TAJ9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Orf178.
GN Orf178.
OS Cafeteria roenbergensis.
OG Eukaryota.
OC Eukaryota; stramenopiles; Bicosoecida; Cafeteria.
OX NCBI_TaxID=33653;
RN [1]
RP SEQUENCE FROM N.A.
RA Burger G.;
RT "The mitochondrial genome of Cafeteria roenbergensis."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF193903; AAF05787.1; -.
KW Mitochondrion.
SQ SEQUENCE 178 AA; 21418 MW; 00F69B2FC8461362 CRC64;

Query Match 93.8%; Score 30; DB 8; Length 178;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 109 GFYRN 113

RESULT 11
Q9VGJ2 PRELIMINARY; PRT; 184 AA.
AC Q9VGJ2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE I-T protein.
GN I-T OR CGI4719.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

```

```

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003692; AAF54687.1; -.
DR FlyBase; FBgn0025821; I-T.
SQ SEQUENCE 184 AA; 21018 MW; 97861392D5B77069 CRC64;

Query Match 93.8%; Score 30; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 31 GFYRN 35

RESULT 12
Q96723 PRELIMINARY; PRT; 184 AA.
AC Q96723;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Inhibitor-T protein.
GN I-T OR INHIBITOR-T OR CGI4719.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN=CANTON S;
RX MEDLINE=99037736; PubMed=9821974;
RA Helps N.R., Vergidou C., Gaskell T., Cohen P.T.W.;
RT "Characterisation of a novel Drosophila melanogaster testis specific
RT PPI inhibitor related to mammalian inhibitor-2: identification of the
RT site of interaction with PPI."
RL FEBS Lett. 438:131-136(1998).

```

```
DR EMBL; AJ006867; CAA07278.1; -.
DR FlyBase; FBgn0025821; I-t.
FT VARIANT 21 21 A -> T.
FT VARIANT 183 183 D -> G.
SQ SEQUENCE 184 AA; 20988 MW; D697139E7271A7B2 CRC64;

Query Match 93.8%; Score 30; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 31 GFYRN 35

RESULT 13
Q8L803 PRELIMINARY; PRT; 191 AA.
AC Q8L803;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative plastid ribosomal protein CL9.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Lu Z.-X., Laroche A., Gaudet D.;
RT "Triticum aestivum putative plastid ribosomal protein CL9 gene.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV123421; RA92711.1; -
DR InterPro; IPR000244; Ribosomal_L9.
DR Pfam; PF03948; Ribosomal_L9_C; 1.
DR Pfam; PF01281; Ribosomal_L9_N; 1.
DR TIGRFAMs; TIGR00158; L9; 1.
DR PROSITE; PS00651; RIBOSOMAL_L9; 1.
KW Ribosomal protein.
SQ SEQUENCE 191 AA; 21582 MW; DB6D7E72B812A34 CRC64;

Query Match 93.8%; Score 30; DB 10; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 67 GFYRN 71

RESULT 14
Q8TNS2 PRELIMINARY; PRT; 216 AA.
AC Q8TNS2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Trk potassium uptake system protein.
GN TRKA OR MA2210.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArelano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
```

```
RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010907; RA055606.1; -.
DR InterPro; IPR003148; TrkA_N.
DR Pfam; PF02254; TrkA-N; 1.
KW Complete proteome.
SQ SEQUENCE 216 AA; 24016 MW; 9E64363A9CE9FBF9 CRC64;
```

Query Match 93.8%; Score 30; DB 17; Length 216;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 175 GFYRN 179

```
RESULT 15
Q8DIE9 PRELIMINARY; PRT; 222 AA.
AC Q8DIE9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Endonuclease III.
GN TLL1641.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005374; BAC09193.1; -.
KW Complete proteome.
SQ SEQUENCE 222 AA; 25106 MW; 1BEEA1B5AA4ED777 CRC64;
```

Query Match 93.8%; Score 30; DB 16; Length 222;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
DB 88 GFYRN 92

Search completed: December 8, 2003, 09:18:39
Job time : 9.35353 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 2.89899 Seconds
(without alignments)
232.212 Million cell updates/sec

Title: US-09-498-556C-79
Perfect score: 32
Sequence: 1 XXGFYRN.7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	93.8	77	B85684	unknown protein en
2	30	93.8	134	S76180	hypothetical prote
3	30	93.8	172	B87703	conserved hypothet
4	30	93.8	175	S73496	hypoxanthine-guani
5	30	93.8	175	F64250	hypoxanthine-guani
6	30	93.8	223	AC2302	endonuclease III [
7	30	93.8	226	C70790	probable Endonucle
8	30	93.8	237	G84299	hypothetical prote
9	30	93.8	361	AG0169	phosphoserine tran
10	30	93.8	362	C82572	phosphoserine amin
11	30	93.8	408	T25524	hypothetical prote
12	30	93.8	415	T46716	hypothetical prote
13	30	93.8	465	A96553	probable myrosinas
14	30	93.8	481	S50053	tryptophan-tRNA li
15	30	93.8	533	1 YRMSCS	monophenol monooxy
16	30	93.8	747	T16274	hypothetical prote
17	30	93.8	824	S40937	hypothetical prote
18	30	93.8	900	S70630	xeroderma pigmento
19	30	93.8	940	S44345	AcrB/AcrD/AcrF fam
20	30	93.8	1041	C87645	probable transport
21	27	84.4	36	H70251	hypothetical prote
22	27	84.4	60	F70187	ribosomal protein
23	27	84.4	199	T39498	hypothetical prote
24	27	84.4	201	A86636	purine-nucleoside
25	27	84.4	234	A84993	hypothetical prote
26	27	84.4	258	T22092	conserved hypothet
27	27	84.4	275	A69413	hypothetical prote
28	27	84.4	299	F64491	hypothetical prote
29	27	84.4			

30	27	84.4	318	2	B36972	aad 5'-region hypo
31	27	84.4	318	2	S33433	hypothetical prote
32	27	84.4	340	2	D69414	hypothetical prote
33	27	84.4	340	2	C75004	hypothetical prote
34	27	84.4	353	2	E64376	endonuclease III -
35	27	84.4	425	2	F69193	acetyl-CoA synthet
36	27	84.4	426	2	F96994	uncharacterized co
37	27	84.4	461	2	E95152	v-type sodium ATP
38	27	84.4	463	2	H85745	unknown protein en
39	27	84.4	463	2	C90873	hypothetical prote
40	27	84.4	465	2	D64881	ycjX protein - Esc
41	27	84.4	465	2	D90866	probable enzyme (i
42	27	84.4	465	2	E85752	probable EC 2.1 en
43	27	84.4	465	2	AC0659	probable ATP-bind
44	27	84.4	471	1	A41706	tryptophan-tRNA li
45	27	84.4	471	1	S62590	peptidyl-prolyl ci

ALIGNMENTS

RESULT 1

B85684
unknown protein encoded by prophage CP-933C [imported] - Escherichia coli (strain O157:H7)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: B85684
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85684
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-77 <STO>
A:Cross-references: GB:AE005174; NID:G12514756; PIDN:AAG55934.1; GSPDB:GN00145; UWGP:Z18
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z1836

Query Match 93.8%; Score 30; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
|||||
DB 15 GFYRN 19

RESULT 2

S76180
hypothetical protein - Synecocystis sp. (strain PCC 6803)
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S76180
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.,
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasudr
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76180
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-134 <KAN>
A:Cross-references: EMBL:D90914; GB:AB001339; NID:G1653477; PIDN:BAAL8439.1; PID:dl0191
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 93.8%; Score 30; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
|||||
Db 27 GFYRN 31

RESULT 3
B87703
conserved hypothetical protein CC3660 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C/Accession: B87703
R;Niernan, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: B87703
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-172 <STO>
A/Cross-references: GB:AE005673; NID:gl3425418; PIDN:AAK25622.1; GSPDB:GN00148
C/Genetics:
A/Gene: CC3660

Query Match 93.8%; Score 30; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
|||||
Db 73 GFYRN 77

RESULT 4
S73496
hypoxanthine-guanine phosphoribosyltransferase hpt - Mycoplasma pneumoniae (strain ATCC
N/A; Alternate names: hypothetical protein K05_orf175
C/Species: Mycoplasma pneumoniae
A/Variety: ATCC 29342
C/Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
C/Accession: S73496
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A/Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A/Reference number: S73327; MUID:97105885; PMID:8948633
A/Accession: S73496
A/Status: Preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-175 <HIM>
A/Cross-references: EMBL:AE000018; GB:U00089; NID:gl673827; PIDN:AAB95818.1; PID:gl67383
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C/Genetics:
A/Gene: hpt
A/Genetic code: SGC3
C/Superfamily: hypoxanthine phosphoribosyltransferase

Query Match 93.8%; Score 30; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
|||||
Db 159 GFYRN 163

RESULT 5
F64250
hypoxanthine-guanine phosphoribosyltransferase - Mycoplasma genitalium
C/Species: Mycoplasma genitalium
C/Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 07-Dec-1999
C/Accession: F64250

R;Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.;
C.A.; Venter, J.C.
Science 270, 397-403, 1995
A/Title: The minimal gene complement of Mycoplasma genitalium.
A/Reference number: A64200; MUID:96026346; PMID:7569993
A/Accession: F64250
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-175 <TIGR>
A/Cross-references: GB:U39733; GB:L43967; NID:gl046177; PIDN:AAB01648.1; PID:gl046178; T:
A/Experimental source: strain G-37
C/Genetics:
A/Genetic code: SGC3
C/Superfamily: hypoxanthine phosphoribosyltransferase

Query Match 93.8%; Score 30; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
|||||
Db 159 GFYRN 163

RESULT 6
AC2302
endonuclease III [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C/Accession: AC2302
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AC2302
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-223 <KUR>
A/Cross-references: GB:BA000019; PIDN:BA075669.1; PID:gl7133104; GSPDB:GN00179
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: nth
C/Superfamily: apurinic/aprymidinic endonuclease III

Query Match 93.8%; Score 30; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
|||||
Db 89 GFYRN 93

RESULT 7
C70790
probable Endonuclease III - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C/Accession: C70790
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: C70790
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-226 <COL>

A;Cross-references: GB:AL022121; GB:AL123456; NID:g3261559; PIDN:CAA17996.1; PID:g296009
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: nth
C;Superfamily: apurinic/apyrimidinic endonuclease III
C;Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein
F;179,186,189,195/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 93.8%; Score 30; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 72 GFYRN 76

RESULT 8
G84299
hypothetical protein Vng1457c [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: G84299
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebbhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: G84299
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <STO>
A;Cross-references: GB:AE004437; NID:g10580953; PIDN:AAG19763.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG1457C

Query Match 93.8%; Score 30; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 191 GFYRN 195

RESULT 9
AG0169
phosphoserine transaminase (EC 2.6.1.52) [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C;Accession: AG0169
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AG0169
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-361 <KUR>
A;Cross-references: GB:ALS90842; PIDN:CAC90218.1; PID:g15979438; GSPDB:GN00175
C;Genetics:
A;Gene: serC
C;Superfamily: phosphoserine aminotransferase
C;Keywords: aminotransferase

Query Match 93.8%; Score 30; DB 2; Length 361;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 283 GFYRN 287

RESULT 10
C82572
phosphoserine aminotransferase XF2326 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: C82572
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82572
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-362 <SIM>
A;Cross-references: GB:AE004043; GB:AE003849; NID:g9107486; PIDN:AAF85125.1; GSPDB:GN001
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, I
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, I
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawas
A;Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuchioka, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; V
C;Genetics:
A;Gene: XF2326
C;Superfamily: phosphoserine aminotransferase

Query Match 93.8%; Score 30; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GFYRN 7
Db 284 GFYRN 288

RESULT 11
T25524
hypothetical protein C06A5.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T25524
R;Davidson, S.; Wohldmann, P.
submitted to the EMBL Data Library, April 1997
A;Description: The sequence of C. elegans cosmid C06A5.
A;Reference number: T25524
A;Accession: T25524
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-408 <DAV>
A;Cross-references: EMBL:U97193; PIDN:AAB52442.1; GSPDB:GN00019; CESP:C06A5.9
A;Experimental source: strain Bristol N2; clone C06A5
C;Genetics:
A;Gene: CESP:C06A5.9
A;Map position: 1
A;Introns: 3/1; 27/3; 86/1; 154/1; 201/3; 240/3; 268/3; 297/3

Query Match 93.8%; Score 30; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7

|||||

Db 229 GFYRN 233

RESULT 12

T46716

Hypothetical protein L4326.05 [imported] - Leishmania major

C/Species: Leishmania major

C/Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000

C/Accession: T46716

R/Volckaert, G.; Ivens, A.C.; Lawson, D.; Quail, M.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, December 1999

A/Reference number: Z23137

A/Accession: T46716

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-415 <VOL>

A/Cross-references: EMBL:AL121861; PIDN:CAB58381.1

A/Experimental source: strain Friedlin

C/Genetics:

A/Note: L4326.05

Query Match

93.8%; Score 30; DB 2; Length 415;

Best Local Similarity

100.0%; Pred. No. 54;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7

|||||

Db 325 GFYRN 329

RESULT 13

A96553

probable myrosinase precursor 53323-50499 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C/Accession: A96553

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: A96553

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-465 <STO>

A/Cross-references: GB:AE005173; NID:gl0092358; PIDN:RAG12767.1; GSPDB:GN00141

C/Genetics:

A/Map position: 1

C/Superfamily: Agrobacterium beta-glucosidase

Query Match

93.8%; Score 30; DB 2; Length 465;

Best Local Similarity

100.0%; Pred. No. 60;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7

|||||

Db 331 GFYRN 335

RESULT 14

S50053

tryptophan-tRNA ligase (EC 6.1.1.2) alpha-2 chain - mouse

C/Species: Mus musculus (house mouse)

C/Date: 07-May-1995 #sequence_revision 10-Nov-1995 #text_change 03-Jun-2002
 C/Accession: S50053; S50052; I49391; S31461; S31462
 R/Fajot, B.; Sarger, C.; Bonnet, J.; Garret, M.

J. Mol. Biol. 242, 599-603, 1994

A/Title: An alternative splicing modifies the C-terminal end of tryptophanyl-tRNA synthetase

A/Reference number: S50052; MUID:95018226; PMID:7932716

A/Accession: S50053

A/Molecule type: mRNA

A/Residues: 1-481 <PAJ>

A/Cross-references: EMBL:X69657; NID:G55437; PIDN:CAA49348.1; PID:G55438

A/Genetics: LSF

A/Note: Intron position was determined by sequencing of genomic DNA

A/Accession: S50052

A/Molecule type: mRNA

A/Residues: 1-475 <PAW>

A/Cross-references: EMBL:X69656; NID:G55435; PIDN:CAA49347.1; PID:G55436

A/Genetics: SSF

R/Kisselev, L.L.

Biochimie 75, 1027-1039, 1993

A/Title: Mammalian tryptophanyl-tRNA synthetases.

A/Reference number: I49391; MUID:94257729; PMID:7515282

A/Accession: I49391

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-481 <RES>

A/Cross-references: EMBL:X69657; NID:G55437; PIDN:CAA49348.1; PID:G55438

C/Genetics: <LSF>

A/Gene: WRS

A/Introns: 475/2

A/Note: the list of introns may be incomplete; clone W13

C/Genetics: <SSP>

A/Gene: WRS

A/Note: clone S5

C/Superfamily: mammalian tryptophan-tRNA ligase; amino acid-tRNA ligase repeat homology
 C/Keywords: alternative splicing; aminoacyl-tRNA synthetase; ligase; protein biosynthesis
 F;1-481/Product: tryptophan-tRNA ligase alpha-2 chain long splice form #status predicted
 F;1-475/Product: tryptophan-tRNA ligase alpha-2 chain short splice form #status predicted
 F;23-68/Domain: amino acid-tRNA ligase repeat homology <ATL>

Query Match 93.8%; Score 30; DB 2; Length 481;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GFYRN 7

|||||

Db 250 GFYRN 254

RESULT 15

YRMSCS

monophenol monooxygenase (EC 1.14.18.1) precursor [validated] - mouse

N/Alternate names: cresolase; monophenol oxidase; phenolase; tyrosinase

C/Species: Mus musculus (house mouse)

C/Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 20-Apr-2000

C/Accession: A27711; A60778; A32429; B32429; S01170; S02278; S15753; I49736

R/Kwon, B.S.; Wakulchik, M.; Haq, A.K.; Halaban, R.; Kestler, D.

Biochem. Biophys. Res. Commun. 153, 1301-1309, 1988

A/Title: Sequence analysis of mouse tyrosinase cDNA and the effect of melanotropin on its

A/Reference number: A27711; MUID:88268910; PMID:3134020

A/Accession: A27711

A/Molecule type: mRNA

A/Residues: 1-533 <KWO>

A/Cross-references: GB:M20234; NID:G202247; PIDN:AAA40516.1; PID:G202248

A/Experimental source: Cloudman S-91 melanoma cells

R/Kwon, B.S.; Haq, A.K.; Wakulchik, M.; Kestler, D.; Barton, D.E.; Francke, U.; Lamoreux,

J. Invest. Dermatol. 93, 589-594, 1989

A/Title: Isolation, chromosomal mapping, and expression of the mouse tyrosinase gene.

A/Reference number: A60778; MUID:90010220; PMID:2507645

A/Accession: A60778

A/Status: translation not shown

A/Molecule type: DNA

A/Residues: 1-273 <KW2>

A/Experimental source: BALB/c

R;Terao, M.; Tabe, L.; Garattini, E.; Sartori, D.; Studer, M.; Mintz, B.
 Biochem. Biophys. Res. Commun. 159, 848-853, 1989
 A:Title: Isolation and characterization of variant cDNAs encoding mouse tyrosinase.
 A:Reference number: A32429; MUID:89193679; PMID:2494997
 A:Accession: A32429
 A:Molecule type: mRNA
 A:Residues: 1-102, 'C', 104-345, 'G', 347-533 <TER>
 A:Cross-references: GB:M24560; NID:G202249; PIDN:AAA40517.1; PID:G202250
 A:Accession: B32429
 A:Molecule type: mRNA
 A:Residues: 1-77, 155-345, 'G', 347-533 <TE2>
 A:Cross-references: GB:M24560
 A:Experimental source: B16 melanoma cells
 R;Mueller, G.; Ruppert, S.; Schmid, E.; Schuetz, G.
 EMBO J. 7, 2723-2730, 1988
 A:Title: Functional analysis of alternatively spliced tyrosinase gene transcripts.
 A:Reference number: S01170; MUID:89030636; PMID:3141148
 A:Accession: S01170
 A:Molecule type: mRNA
 A:Residues: 1-102, 'C', 104-263, 'I', 265-345, 'G', 347-533 <MUE>
 A:Cross-references: GB:X12782; NID:G55061; PIDN:CAA31273.1; PID:G55062
 R;Yamamoto, H.; Takeuchi, S.; Kudo, T.; Makino, K.; Nakata, A.; Shinoda, T.; Takeuchi, T.
 Jpn. J. Genet. 62, 271-274, 1987
 A:Title: Cloning and sequencing of mouse tyrosinase cDNA.
 A:Reference number: S02278
 A:Accession: S02278
 A:Molecule type: mRNA
 A:Residues: 1-102, 'C', 104-263, 'I', 265-345, 'G', 347-448 <YAM>
 A:Cross-references: EMBL:X12782
 A:Note: part of this sequence was confirmed by protein sequencing
 R;Shibahara, S.; Okinaga, S.; Tomita, Y.; Takeda, A.; Yamamoto, H.; Sato, M.; Takeuchi, R.
 Eur. J. Biochem. 189, 455-461, 1990
 A:Title: A point mutation in the tyrosinase gene of BALB/c albino mouse causing the cyst
 A:Reference number: S15753; MUID:90249393; PMID:2110899
 A:Accession: S15753
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-13 <SHI>
 A:Cross-references: EMBL:X51743; NID:G55057; PIDN:CAA36033.1; PID:G55058
 A:Experimental source: strain BALB/c
 R;Kwon, B.S.; Halaban, R.; Chintaneni, C.
 Biochem. Biophys. Res. Commun. 161, 252-260, 1989
 A:Title: Molecular basis of mouse Himalayan mutation.
 A:Reference number: I49736; MUID:89273644; PMID:2567165
 A:Accession: I49736
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: mRNA
 A:Residues: 1-39, 'I', 41-102, 'C', 104-196, 'Q', 198-345, 'G', 347-419, 'R', 421-533 <RES>
 A:Cross-references: GB:M26729; NID:G193845; PIDN:AAA37806.1; PID:G309296
 C:Comment: This cell-specific oxidase is a glycoprotein containing two Cu per enzyme; it
 reactions in the formation of pigments such as melanins and other polyphenolic compound
 C:Genetics:
 A:Gene: Tyrl
 A:Map position: 7
 C:Superfamily: monophenol monooxygenase
 C:Keywords: albinism; alternative splicing; copper; glycoprotein; melanin biosynthesis;
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-533/Product: monophenol monooxygenase #status predicted <MAT>
 F:474-497/Domain: transmembrane #status predicted <TM>
 F:86,111,161,230,337,371/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.8%; Score 30; DB 1; Length 533;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GYRN 7
 |||||
 Db 458 GYRN 462

Search completed: December 8, 2003, 09:19:32
 Job time : 4.89899 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:31 ; Search time 1.85859 Seconds
(without alignments)
91.060 Million cell updates/sec

Title: US-09-498-556C-85

Perfect score: 4

Sequence: 1 XXXX 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*

5: /cgn2_6/ptodata/1/1aa/PTCUTUS_COMB.pep:*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	1	1	US-07-820-154A-12
2	0	0.0	1	1	US-07-791-213D-24
3	0	0.0	1	1	US-07-791-213D-40
4	0	0.0	1	1	US-08-174-365A-57
5	0	0.0	1	1	US-07-789-913-23
6	0	0.0	1	1	US-07-789-913-25
7	0	0.0	1	1	US-08-049-794-23
8	0	0.0	1	1	US-08-049-794-25
9	0	0.0	1	1	US-08-433-037-12
10	0	0.0	1	1	US-08-448-606-4
11	0	0.0	1	1	US-07-869-933-16
12	0	0.0	1	1	US-08-293-150A-24
13	0	0.0	1	1	US-08-293-150A-40
14	0	0.0	1	1	US-08-496-847-23
15	0	0.0	1	1	US-08-496-847-25
16	0	0.0	1	2	US-08-742-774-23
17	0	0.0	1	2	US-08-742-774-25
18	0	0.0	1	2	US-08-675-354-23
19	0	0.0	1	2	US-08-675-354-25
20	0	0.0	1	2	US-08-097-554A-12
21	0	0.0	1	2	US-08-965-918-23
22	0	0.0	1	2	US-08-965-918-25
23	0	0.0	1	2	US-09-138-439-23
24	0	0.0	1	2	US-09-138-439-25
25	0	0.0	1	3	US-08-480-640A-12
26	0	0.0	1	3	US-08-613-400A-23
27	0	0.0	1	3	US-08-613-400A-25

28	0	0.0	1	3	US-08-801-092-10	Sequence 10, Appl
29	0	0.0	1	3	US-08-801-092-17	Sequence 17, Appl
30	0	0.0	1	3	US-08-801-092-24	Sequence 24, Appl
31	0	0.0	1	3	US-08-801-092-31	Sequence 31, Appl
32	0	0.0	1	3	US-08-801-092-38	Sequence 38, Appl
33	0	0.0	1	3	US-08-801-092-45	Sequence 45, Appl
34	0	0.0	1	3	US-09-298-017-23	Sequence 23, Appl
35	0	0.0	1	3	US-09-298-017-25	Sequence 25, Appl
36	0	0.0	1	3	US-08-295-802-12	Sequence 12, Appl
37	0	0.0	1	3	US-09-392-979A-23	Sequence 23, Appl
38	0	0.0	1	3	US-09-392-979A-25	Sequence 25, Appl
39	0	0.0	1	3	US-09-103-663-16	Sequence 16, Appl
40	0	0.0	1	3	US-08-488-237A-12	Sequence 12, Appl
41	0	0.0	1	3	US-09-117-927-5	Sequence 5, Appl
42	0	0.0	1	4	US-08-375-992A-12	Sequence 12, Appl
43	0	0.0	1	4	US-09-315-113-10	Sequence 10, Appl
44	0	0.0	1	4	US-09-315-113-17	Sequence 17, Appl
45	0	0.0	1	4	US-09-315-113-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-07-820-154A-12
; Sequence 12, Application US/07820154A
; Patent No. 5382425
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D., Mark D
; APPLICANT: Junker M.S., David E
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/820,154A
; APPLICATION NUMBER: US/07/820,154A
; FILING DATE: 19920113
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-820-154A-12

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 X 1

Db 1 V 1

RESULT 2

US-07-791-213D-24

Sequence 24, Application US/07791213D
Patent No. 5409895
GENERAL INFORMATION:
APPLICANT: MORISHITA, Hideaki
APPLICANT: KANAMORI, Toshinori
APPLICANT: NOBUHARA, Masahiro
TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
TITLE OF INVENTION: TREATING USING THE SAME
NUMBER OF SEQUENCES: 108
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07791,213D
FILING DATE: 13-NOV-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-306745
FILING DATE: 13-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 029650-032
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-791-213D-24

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 3
US-07-791-213D-40
Sequence 40, Application US/07791213D
Patent No. 5409895
GENERAL INFORMATION:
APPLICANT: MORISHITA, Hideaki
APPLICANT: KANAMORI, Toshinori
APPLICANT: NOBUHARA, Masahiro
TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
TITLE OF INVENTION: TREATING USING THE SAME
NUMBER OF SEQUENCES: 108
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States

ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07791,213D
FILING DATE: 13-NOV-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-306745
FILING DATE: 13-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 029650-032
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-791-213D-40

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 4
US-08-174-365A-57
Sequence 57, Application US/08174365A
Patent No. 5478809
GENERAL INFORMATION:
APPLICANT: Seichi TANIDA et al.
TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/174,365A
FILING DATE: December 28, 1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:

LENGTH: 1 amino acid
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: modified site
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: /note = "Xaa is modified amino acid as
OTHER INFORMATION: described in specification"
US-08-174-365A-57

Query Match 0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 X 1

Db 1 X 1

RESULT 5

US-07-789-913-23

Sequence 23, Application US/07789913
Patent No. 5559095
GENERAL INFORMATION:
APPLICANT: Miljanich, George P.
APPLICANT: Bowersox, Stephen S.
APPLICANT: Fox, James A.
APPLICANT: Valentino, Karen L.
APPLICANT: Bitner, Robert S.
APPLICANT: Yamashiro, Donald H.
TITLE OF INVENTION: Delayed Treatment Method of Reducing
TITLE OF INVENTION: Ischemia-Related Neuronal Damage
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Peter Dehlinger
STREET: 350 Cambridge Avenue, Suite 300
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 19911112
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/561,766
FILING DATE: 02-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/440,094
FILING DATE: 22-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Stratford, Carol A.
REGISTRATION NUMBER: 34,444
REFERENCE/DOCKET NUMBER: 5865-0005.30
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: AMINO ACID
TOPOLOGY: both
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO

ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: peptide fragment used in the claims
US-07-789-913-23

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 C 1

RESULT 6

US-07-789-913-25
Sequence 25, Application US/07789913
Patent No. 5559095
GENERAL INFORMATION:
APPLICANT: Miljanich, George P.
APPLICANT: Bowersox, Stephen S.
APPLICANT: Fox, James A.
APPLICANT: Valentino, Karen L.
APPLICANT: Bitner, Robert S.
APPLICANT: Yamashiro, Donald H.
TITLE OF INVENTION: Delayed Treatment Method of Reducing
TITLE OF INVENTION: Ischemia-Related Neuronal Damage
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Peter Dehlinger
STREET: 350 Cambridge Avenue, Suite 300
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/789,913
FILING DATE: 19911112
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/561,766
FILING DATE: 02-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/440,094
FILING DATE: 22-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Stratford, Carol A.
REGISTRATION NUMBER: 34,444
REFERENCE/DOCKET NUMBER: 5865-0005.30
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: AMINO ACID
TOPOLOGY: both
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: peptide fragment used in the claims
US-07-789-913-25

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 R 1

RESULT 7

US-08-049-794-23

Sequence 23, Application US/08049794

Patent No. 5587454

GENERAL INFORMATION:

APPLICANT: JUSTICE, ALAN

APPLICANT: SINGH, TEJINDER

APPLICANT: GOHIL, KISHOR C

APPLICANT: VALENTINO, KAREN L

APPLICANT: MILJANICH, GEORGE P

TITLE OF INVENTION: METHODS OF PRODUCING ANALGESIA AND

ENHANCING OPIATE ANALGESIA

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Law Offices of Peter Dehlinger

STREET: 350 Cambridge Avenue, Suite 300

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/049,794

FILING DATE: 19930415

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/814,759

FILING DATE: 30-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: Stratford, Carol A.

REGISTRATION NUMBER: 34,444

REFERENCE/DOCKET NUMBER: 5865-0009.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 1 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT, PAGE

INDIVIDUAL ISOLATE: 32

US-08-049-794-23

Query Match 0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 C 1

RESULT 8

US-08-049-794-25

Sequence 25, Application US/08049794

Patent No. 5587454

GENERAL INFORMATION:

APPLICANT: JUSTICE, ALAN

APPLICANT: SINGH, TEJINDER

APPLICANT: GOHIL, KISHOR C

APPLICANT: VALENTINO, KAREN L

TITLE OF INVENTION: METHODS OF PRODUCING ANALGESIA AND

ENHANCING OPIATE ANALGESIA

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: Law Offices of Peter Dehlinger

STREET: 350 Cambridge Avenue, Suite 300

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/049,794

FILING DATE: 19930415

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/814,759

FILING DATE: 30-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: Stratford, Carol A.

REGISTRATION NUMBER: 34,444

REFERENCE/DOCKET NUMBER: 5865-0009.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 1 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT, PAGE

INDIVIDUAL ISOLATE: 32

US-08-049-794-23

Query Match 0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 C 1

RESULT 9

US-08-433-037-12

Sequence 12, Application US/08433037

Patent No. 5707828

GENERAL INFORMATION:

APPLICANT: Sreekrishna, Kotikanyadan

APPLICANT: Barr, Kathryn A.

APPLICANT: Brierley, Russell A.

APPLICANT: Thill, Gregory P.

APPLICANT: Tschoop, Juerg F.

TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: U.S.A.

ZIP: 11530-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

US-08-049-794-25

Query Match 0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 R 1

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/433,037
FILING DATE: 03-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 91082
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-12

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 L 1

RESULT 10
US-08-448-606-4
Sequence 4, Application US/08448606
Patent No. 5721114
GENERAL INFORMATION:
APPLICANT: Abrahams n, Lars
APPLICANT: Holmgren, Erik
APPLICANT: Kalder n, Christina
APPLICANT: Lake, Mats
APPLICANT: Mikaelsson, sa
APPLICANT: Sejlitz, Torsten
TITLE OF INVENTION: Expression System For Producing
TITLE OF INVENTION: Apolipoprotein AI-M
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pollock, Vande Sande & Priddy
STREET: 1990 M Street, N.W., Suite 800
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,606
FILING DATE: 25-AUG-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SE93/01061
FILING DATE: 09-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9203753-0
FILING DATE: 11-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Amernick, Burton A.
REGISTRATION NUMBER: 24,852
REFERENCE/DOCKET NUMBER: 0151/00121
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)331-7111
TELEFAX: (202) 293-6229
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acid
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-448-606-4

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 Q 1

RESULT 11
US-07-869-933-16
Sequence 16, Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:
APPLICANT: KINET, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
TITLE OF INVENTION: IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869,933
FILING DATE: 19920416
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/154 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-869-933-16

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 M 1

RESULT 12
US-08-293-150A-24
Sequence 24, Application US/08293150A

Patent No. 5792629
GENERAL INFORMATION:
APPLICANT: MORISHITA, Hideaki
APPLICANT: KANAMORI, Toshinori
APPLICANT: NOBUHARA, Masahiro
TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
TITLE OF INVENTION: TREATING USING THE SAME
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/293,150A
FILING DATE: 19-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/791,213
FILING DATE: 13-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-306745
FILING DATE: 13-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M.
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 029650-049
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-293-150A-24

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 13
US-08-293-150A-40
Sequence 40, Application US/08293150A
Patent No. 5792629
GENERAL INFORMATION:
APPLICANT: MORISHITA, Hideaki
APPLICANT: KANAMORI, Toshinori
APPLICANT: NOBUHARA, Masahiro
TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
TITLE OF INVENTION: TREATING USING THE SAME
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: P.O. Box 1404
CITY: Alexandria

STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/293,150A
FILING DATE: 19-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/791,213
FILING DATE: 13-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-306745
FILING DATE: 13-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M.
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 029650-049
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-293-150A-40

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 14
US-08-496-847-23
Sequence 23, Application US/08496847
Patent No. 5795864
GENERAL INFORMATION:
APPLICANT: Ametutz, Gary A.
APPLICANT: Bowersox, Stephen S.
APPLICANT: Gohil, Kishorchandra
APPLICANT: Adriaenssens, Peter I.
APPLICANT: Kristipati, Ramasharma
TITLE OF INVENTION: METHODS AND
TITLE OF INVENTION: FORMULATIONS FOR PREVENTING PROGRESSION OF NEUROPATHIC PAIN
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94306-1546
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/496,847
FILING DATE: 27-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Stratford, Carol A

REGISTRATION NUMBER: 34,444
REFERENCE/DOCKET NUMBER: 5865-0009.31
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT
US-08-496-847-23

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 C 1

RESULT 15

US-08-496-847-25
Sequence 25, Application US/08496847
Patent No. 5795864

GENERAL INFORMATION:

APPLICANT: Amstutz, Gary A.
APPLICANT: Bowersox, Stephen S.
APPLICANT: Gohil, Kishorchandra
APPLICANT: Adriaenssens, Peter I.
APPLICANT: Kristipati, Ramasharma

TITLE OF INVENTION: METHODS AND

FORMULATIONS FOR PREVENTING PROGRESSION OF NEUROPATHIC PAIN

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94306-1546

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/496,847

FILING DATE: 27-JUN-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Stratford, Carol A

REGISTRATION NUMBER: 34,444

REFERENCE/DOCKET NUMBER: 5865-0009.31

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-324-0880

TELEFAX: 650-324-0960

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 1 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT

US-08-496-847-25

Query Match

0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 0.0%; Pred.No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 R 1

Search completed: December 8, 2003, 09:20:29

Job time : 1.85859 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:18:46 ; Search time 3.55556 Seconds
(without alignments)
209.232 Million cell updates/sec

Title: US-09-498-556C-85
Perfect score: 4
Sequence: 1 XXXX 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	1	9 US-09-909-348-4	Sequence 4, Appli
2	0	0.0	1	10 US-09-982-172-3	Sequence 3, Appli
3	0	0.0	1	10 US-09-982-172-4	Sequence 4, Appli
4	0	0.0	1	10 US-09-982-172-9	Sequence 9, Appli
5	0	0.0	1	10 US-09-982-172-11	Sequence 11, Appli
6	0	0.0	1	10 US-09-982-172-19	Sequence 19, Appli
7	0	0.0	1	10 US-09-982-172-31	Sequence 31, Appli
8	0	0.0	1	10 US-09-982-172-35	Sequence 35, Appli
9	0	0.0	1	10 US-09-982-172-37	Sequence 37, Appli
10	0	0.0	1	10 US-09-982-172-46	Sequence 46, Appli
11	0	0.0	1	10 US-09-982-172-69	Sequence 69, Appli
12	0	0.0	1	10 US-09-982-172-80	Sequence 80, Appli
13	0	0.0	1	10 US-09-982-172-81	Sequence 81, Appli
14	0	0.0	1	10 US-09-982-172-83	Sequence 83, Appli
15	0	0.0	1	10 US-09-982-172-86	Sequence 86, Appli

16	0	0.0	1	10 US-09-982-172-93	Sequence 93, Appli
17	0	0.0	1	10 US-09-982-172-95	Sequence 95, Appli
18	0	0.0	1	10 US-09-982-172-106	Sequence 106, Appli
19	0	0.0	1	10 US-09-982-172-112	Sequence 112, Appli
20	0	0.0	1	10 US-09-982-172-120	Sequence 120, Appli
21	0	0.0	1	10 US-09-982-172-126	Sequence 126, Appli
22	0	0.0	1	10 US-09-982-172-148	Sequence 148, Appli
23	0	0.0	1	10 US-09-982-172-149	Sequence 149, Appli
24	0	0.0	1	10 US-09-982-172-155	Sequence 155, Appli
25	0	0.0	1	10 US-09-982-172-160	Sequence 160, Appli
26	0	0.0	1	10 US-09-982-172-172	Sequence 172, Appli
27	0	0.0	1	10 US-09-982-172-173	Sequence 173, Appli
28	0	0.0	1	10 US-09-982-172-175	Sequence 175, Appli
29	0	0.0	1	10 US-09-982-172-189	Sequence 189, Appli
30	0	0.0	1	10 US-09-982-172-190	Sequence 190, Appli
31	0	0.0	1	10 US-09-982-172-191	Sequence 191, Appli
32	0	0.0	1	10 US-09-982-172-195	Sequence 195, Appli
33	0	0.0	1	10 US-09-982-172-200	Sequence 200, Appli
34	0	0.0	1	10 US-09-982-172-211	Sequence 211, Appli
35	0	0.0	1	11 US-09-809-391-395	Sequence 395, Appli
36	0	0.0	1	11 US-09-809-391-611	Sequence 611, Appli
37	0	0.0	1	12 US-10-032-221B-19	Sequence 19, Appli
38	0	0.0	1	12 US-10-299-619-4	Sequence 4, Appli
39	0	0.0	1	12 US-10-061-607A-28	Sequence 28, Appli
40	0	0.0	1	12 US-10-061-607A-29	Sequence 29, Appli
41	0	0.0	1	12 US-10-061-607A-35	Sequence 35, Appli
42	0	0.0	1	12 US-09-882-171-395	Sequence 395, Appli
43	0	0.0	1	12 US-09-882-171-611	Sequence 611, Appli
44	0	0.0	1	12 US-10-104-307-7	Sequence 7, Appli
45	0	0.0	1	12 US-10-104-307-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-909-348-4
; Sequence 4, Application US/0909348
; Patent No. US20020042373A1
; GENERAL INFORMATION:
; APPLICANT: Carney, Darrell H.
; APPLICANT: Crowther, Roger S.
; APPLICANT: Stierberg, Janet
; APPLICANT: Bergmann, John
; TITLE OF INVENTION: Stimulation Of Cartilage Growth With Agonists
; FILE REFERENCE: Of The No. US20020042373A1-Proteolytically Activated Thrombin R
; CURRENT APPLICATION NUMBER: US/09/909,348
; PRIOR FILING DATE: 2001-07-19
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide fragment of Thrombin
; NAME/KEY: VARIANT
; LOCATION: (1)...(14)
; OTHER INFORMATION: Xaa at position six is Glu or Gln
; OTHER INFORMATION: Xaa at position thirteen is Phe, Met, Leu, His or
US-09-909-348-4

Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 V 1

```
RESULT 2
US-09-982-172-3
; Sequence 3, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-3

Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 3
US-09-982-172-4
; Sequence 4, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-4

Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 4
US-09-982-172-9
; Sequence 9, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; FILE REFERENCE: 01/22283
```

```
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-9

Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 5
US-09-982-172-11
; Sequence 11, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-11

Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 6
US-09-982-172-19
; Sequence 19, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-19
```

Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 K 1

RESULT 7

US-09-982-172-31
; Sequence 31, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-31

Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 K 1

RESULT 8

US-09-982-172-35
; Sequence 35, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-35

Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 R 1

RESULT 9

US-09-982-172-37
; Sequence 37, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-37

Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 K 1

RESULT 10

US-09-982-172-46
; Sequence 46, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-46

Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 K 1

RESULT 11

US-09-982-172-69
; Sequence 69, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19

; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-69

Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1
DB 1 K 1

RESULT 12
US-09-982-172-80
; Sequence 80, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 80
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-80

Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1
DB 1 K 1

RESULT 13
US-09-982-172-81
; Sequence 81, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 81
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-81

Query Match 0.0%; Score 0; DB 10; Length 1;

Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1
DB 1 K 1

US-09-982-172-83
; Sequence 83, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 83
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-83

Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1
DB 1 R 1

RESULT 15
US-09-982-172-86
; Sequence 86, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 86
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-86

Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1
DB 1 K 1

Search completed: December 8, 2003, 09:34:03
Job time : 3.55556 secs

Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 R 1

RESULT 14
US-09-982-172-83
; Sequence 83, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 83
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-83

Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 R 1

RESULT 15
US-09-982-172-86
; Sequence 86, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 86
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-86

Query Match 0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 K 1

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 34.7879 Seconds
(without alignments)
232.212 Million cell updates/sec

Title: US-09-498-556C-59

Perfect score: 486
Sequence: 1 KATMOGENEKYDSCGSKC.....VSAEDCLNDMDIYPGTRN 84

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	125	25.7	561	2 T27318	hypothetical prote
2	125	25.7	626	2 T27319	hypothetical prote
3	113	23.3	490	2 T32003	hypothetical prote
4	112	23.0	249	2 T24604	hypothetical prote
5	109.5	22.5	1642	2 T19130	hypothetical prote
6	105.5	21.7	145	2 T15608	hypothetical prote
7	100	20.6	5376	2 T42215	zonadhesin - mouse
8	99	20.4	1036	2 T17405	scavenger receptor
9	98.5	20.3	135	2 T15610	hypothetical prote
10	98	20.2	137	2 T15609	hypothetical prote
11	96.5	19.9	869	1 JC4858	LDL receptor prec
12	95.5	19.7	166	2 H89044	protein B0238.12 f
13	94.5	19.4	98	2 C89046	protein C1068.4 i
14	94.5	19.4	4660	2 T42737	sp330 protein prec
15	92.5	19.0	63	2 S07127	chymotrypsin/elast
16	92.5	19.0	63	2 S08572	chymotrypsin/elast
17	92.5	19.0	4753	1 A47437	LDL-receptor-relat
18	90.5	18.6	195	2 T28803	hypothetical prote
19	87.5	18.0	802	2 T24293	hypothetical prote
20	87.5	18.0	949	2 T24294	hypothetical prote
21	85.5	17.6	863	1 S51789	hypothetical prote
22	85	17.5	2155	2 T30197	VLDL receptor prec
23	84.5	17.4	13288	2 T03099	alpha tectorin - m
24	83	17.1	209	2 T02394	mucin, submaxillar
25	82.5	17.0	873	1 A49729	hypothetical prote
26	82.5	17.0	873	1 ORRBVD	VLDL receptor prec
27	82	16.9	62	2 S35098	trypsin inhibitor
28	82	16.9	267	2 T30007	hypothetical prote
29	81.5	16.8	1810	1 A32230	tenascin precursor

30	81.5	16.8	4135	2 T42629	tenascin-X - bovin
31	80.5	16.6	169	2 T15611	hypothetical prote
32	80.5	16.6	4006	2 T09070	probable tenascin
33	79.5	16.4	2813	1 VWHU	von Willebrand fac
34	79.5	16.4	3002	2 A47221	fibrillin 1 precu
35	79	16.3	548	2 T16642	hypothetical prote
36	79	16.3	647	2 A43902	tenascin - eastern
37	78.5	16.2	798	2 A28193	integrin beta-1 ch
38	78.5	16.2	1160	2 F88369	protein unc-52 f
39	78.5	16.2	1321	2 J80352	mucin MUC5B, trach
40	78.5	16.2	2295	2 C88369	protein unc-52 f
41	78.5	16.2	3375	2 T19821	hypothetical prote
42	78	16.0	305	2 A46476	B cell-associated
43	78	16.0	1107	2 T15884	hypothetical prote
44	78	16.0	1847	2 T18308	probable vitalloge
45	78	16.0	3133	2 S52093	hemocytin - silkw

ALIGNMENTS

RESULT 1

T27318

hypothetical protein Y69H2.3a - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27318

R:McMurray, A.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z20343

A:Accession: T27318

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-561 <WIL>

A:Cross-references: EMBL:Z98877; PIDN:CAB54472.1; GSPDB:GN00023; CESP:Y69H2.3a

A:Experimental source: clone Y69H2

C:Genetics:

A:Gene: CESP:Y69H2.3a

A:Map position: 5

A:Introns: 183/1; 247/1; 288/1; 318/1; 364/1; 388/2; 457/1; 481/2; 521/3

Query Match 25.7%; Score 125; DB 2; Length 561;

Best Local Similarity 36.1%; Pred. NO. 0.00025;

Matches 26; Conservative 10; Mismatches 16; Indels 20; Gaps 3;

QY	5	QCGENEKYDSCGSKCKKCYDGVVEEDDEPNVPCLVRVCHQ-----DCVCEGFPVRN	59
DB	498	RCRSNEKEPCKTCVCSDTKC-----NNEP-----RFCPQVCTGGGCVQSGFFRD	542

QY	60	KDDKCVSAEDCE	71
DB	543	NSGKCVTQNDCE	554

RESULT 2

T27319

hypothetical protein Y69H2.3b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27319

R:McMurray, A.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z20343

A:Accession: T27319

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-626 <WIL>

A:Cross-references: EMBL:Z98877; PIDN:CAB54473.1; GSPDB:GN00023; CESP:Y69H2.3b

A:Experimental source: clone Y69H2

C:Genetics:

A:Gene: CESP:Y69H2.3b

A:Map position: 5

A:Introns: 183/1; 247/1; 312/1; 353/1; 383/1; 429/1; 453/2; 522/1; 546/2; 586/3

Qy	6	CGENENKYDSCGSKCKKCYDGVVEEDDEBPVPCLVRVCHQDVCVCEGFYRNKDDKCV	65
Db	185	CGPNEHFVCGKNICSDTKC-----NEKRMKCPACTFFGCVCLNGFFRDKDKCV	234
Qy	66	SAEDCE 71	
Db	235	TOEED 240	
RESULT 5			
T19130			
hypothetical protein C09F9.2 - Caenorhabditis elegans			
C/Species: Caenorhabditis elegans			
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999			
C/Accession: T19130			
R:Smve, R.			
submitted to the EMBL Data Library, November 1996			
A/Reference number: Z19078			
A/Accession: T19130			
A/Status: preliminary; translated from GB/EMBL/DBDJ			
A/Molecule type: DNA			
A/Residues: 1-1642 <WII>			
A/Cross-references: EMBL:Z81465; PIDN:CAB03861.1; GSPDB:GN00020; CBSP:C09F9.2			
A/Experimental source: clone C09F9			
C/Genetics:			
A/Map position: 2			
A/Introns: 75/3; 103/2; 220/1; 552/1; 1500/3; 1517/1; 1578/3			
Query Match 22.5%; Score 109.5; DB 2; Length 1642;			
Best Local Similarity 33.8%; Pred. No. 0.016;			
Matches 24; Conservative 11; Mismatches 27; Indels 9; Gaps 3			
Qy	5	QCGENKYDSC--GSKCEDKKCYDGVVEEDDEBPVPCLVRVCHQDVCVCEGFYR--NK	60
Db	319	ECGPNQWSECPSSRECHESCDWTHFFETTPNCPNSCGTPR-----CIKKGGFVRMAD	373
Qy	61	DDKCVSAEDCE 71	
Db	374	EDVCVPFFDCD 384	
RESULT 6			
T15608			
hypothetical protein C25E10.7 - Caenorhabditis elegans			
C/Species: Caenorhabditis elegans			
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999			
C/Accession: T15608			
R:Bradshaw, H.			
submitted to the EMBL Data Library, February 1996			
A/Description: The sequence of C. elegans cosmid C25E10.			
A/Reference number: Z18376			
A/Accession: T15608			
A/Status: preliminary; translated from GB/EMBL/DBDJ			
A/Molecule type: DNA			
A/Residues: 1-145 <BRA>			
A/Cross-references: EMBL:U50311; NID:g1226295; PID:g1226302; PIDN:AAA92312.1;			
A/Experimental source: strain Bristol N2			
C/Genetics:			
A/Gene: CBSP:C25E10.7			
A/Introns: 39/3			
Query Match 21.7%; Score 105.5; DB 2; Length 145;			
Best Local Similarity 31.8%; Pred. No. 0.0052;			
Matches 21; Conservative 14; Mismatches 20; Indels 11; Gaps 3			
Qy	5	QCGENKYDSCGSKCKKCYDGVVEEDDEBPVPCLVRVCHQDVCVCEGFYRNKDDK	64
Db	17	ECGKNQKRVACG-YDCEPQCGFD-----PTV-CSLECKPNACVKDGYVRNTKND	65
Qy	65	VSAREDC 70	
Db	66	VRELEC 71	

A;Molecule type: mRNA
A;Residues: 1-5376 <GAG>
A;Cross-references: EMBL:U97069; NID:G3327420; PID:G3327421; PIDN:AAC26680.1
C;Genetics:
A;Gene: Zan
A;Map position: 5
C;Function:
A;Description: functions in multiple cell adhesion processes
A;Note: found exclusively on the apical region of the sperm head
C;Keywords: cell adhesion

Query Match 20.3%; Score 98.5; DB 2; Length 135;
Best Local Similarity 32.0%; Pred.No. 0.022;
Matches 24; Conservative 11; Mismatches 19; Indels 2

Qy 1 KATMOGENEKYDSCGSKCEDKKCKYDGVVEEDDEPN-----VPCLVRVCHQDCVCEEG 55

D6 75 KETTKCPENETFFCGT-ACEATC-----EKNPTVTKQCI VNVCO-----CSKG 119

QY	56	FYRNKDDKCVSAEDC	70
		:	
Db	120	FVRH-GLRCIDKDC	133

```

RESULT 10
T15609
hypothetical protein C25E10.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #te
C:Accession: T15609
R:Bradshaw, H.
submitted to the EMBL Data Library, February 1996
A:Description: The sequence of C. elegans cosmid C25E1
A:Reference number: Z19376.

```

A/Accession: F113609
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Molecule type: DNA
A/Residues: 1-137

A/Cross-references: EMBL:U50311; NID:G1262695; PID:G1262695
A/Experimental source: strain Bristol N2
C/Genetics:
A/Gene: CESP:C25E10.8
A/Introns: 1/3; 19/1; 42/3

Query Match	20.2%;	Score 98;	DB 2;	Length 137;
Best Local Similarity	31.8%;	Pred. No. 0.025;		
Matches	28;	Conservative	9;	Mismatches 33; Indels 18; Gaps 5;
Qy	2	ATWQGENEKYDSCGSKCDKKCKYDGVSEEDDEBPWCLVRVCHQD-CVCEEGFYRNK	60	
Db	17	ANKTCGANEEMVAC-HDHCTQCGY-----TPKV-CLSAQCIENACDCKGQVRNS	65	
Qy	61	DKCVSAEDCEL-----DNMDFIYPGT	82	
Db	66	LGKCVDISTCTKTSKCPENETFFRCGT	93	

RESULT 11 .
JC4858

J4858
 VitD receptor precursor - African clawed frog
 N;Alternate names: very low density lipoprotein receptor; vitellogenin receptor
 C;Species: Xenopus laevis (African clawed frog)
 C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 16-Jun-2000
 C;Accession: J4858
 R;Okabayashi, K.; Shoji, H.; Nakamura, T.; Hashimoto, O.; Asashima, M.; Sugino, H.
 Biochem. Biophys. Res. Commun. 224, 406-413, 1996
 A;Title: cDNA cloning and expression of the Xenopus laevis vitellogenin receptor.
 A;Reference number: J4858; MUID:96295501; PMID:8702402
 A;Accession: J4858

A:Molecule type: mRNA
A:Residues: 1-869 <OKA>
A:Cross-references: GB:AB006906; MID:g2366772; PIDN:BAA22145.1; PID:g2366773
C:Comment: This receptor mediates incorporation of vitellogenin into oocytes.
C:Superfamily: LDL receptor; EGF homology; LDL receptor ligand-binding repeat homology;
C:Keywords: duplication; fatty acid metabolism; glycoprotein; receptor; transmembrane p
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-869/Product: VLDL receptor #status predicted <MAT>
F:27-793/Domain: extracellular #status predicted <EXT>
F:32-66/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F:71-107/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F:112-148/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F:153-187/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F:192-228/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F:238-272/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F:277-311/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F:317-354/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F:359-393/Domain: EGF homology <EG1>
F:399-433/Domain: EGF homology <EG2>
F:440-479/Domain: LDL receptor WTD-containing repeat homology <YW1>
F:480-525/Domain: LDL receptor WTD-containing repeat homology <YW2>
F:526-568/Domain: LDL receptor WTD-containing repeat homology <YW3>
F:569-612/Domain: LDL receptor WTD-containing repeat homology <YW4>
F:613-655/Domain: LDL receptor WTD-containing repeat homology <YW5>
F:656-698/Domain: LDL receptor WTD-containing repeat homology <YW6>
F:707-749/Domain: EGF homology <EG3>
F:794-815/Domain: transmembrane #status predicted <TM>
F:816-869/Domain: intracellular #status predicted <CYT>
F:830-834/Region: coated-pit mediated internalization signal
F:150,201,777,786/Binding site: carbohydrate (Aan) (covalent) #status predicted
F:359-370,366-379,381-393,399-409,405-418,420-433,707-720,716-735,737-749/Disulfide bond

Query Match 19.9%; Score 96.5; DB 1; Length 869;
Best Local Similarity 35.4%; Pred. No. 0.16;
Matches 28; Conservative 8; Mismatches 38; Indels 5; Gaps 4;

Qy 2 ATMOCGENEKYDSCGSKCEDKKCKYDGVVEEDDEPNVPLVVRVCHOD-CVCEGFYRN 59
Db 234 APQRCANEM--PCGSECTHKRWKRCDDADCKRDSBEINCPSTCTCPDQFKEDGNCIH 291

Qy 60 KDKCVSAEDCELDNMDFI 78
Db 292 GSRQCDGVDRDC-LDGTDEI 309

RESULT 12
H89044
protein B0238.12 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: H89044
R:anonymous, The C. elegans Sequencing Consortium.
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: H89044
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-166 <STO>
A:Cross-references: GB:chr_V; PIDN:AAB65990.1; PID:g2315490; GSPDB:GN00023; CESP:B0238.1
C:Genetics:
A:Map position: 5

Query Match 19.7%; Score 95.5; DB 2; Length 166;
Best Local Similarity 31.4%; Pred. No. 0.049;
Matches 22; Conservative 9; Mismatches 28; Indels 11; Gaps 4;

Qy 1 KATMOCCGENEKYDSCGSKCEDKKCKYDGVVEEDDEPNVPLVVRVCHODCVCEGFYRNK 60
Db 108 QSPQRCGRNETFRTCGS-SCEPSC-----TTFRPQACTMQCIVNVCO-----CSEGFVRGP 157

Qy 61 DDKCVSAEDC 70
Db 158 SG-CVRQRDC 166

RESULT 13
C89046
protein C10G8.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C89046
R:anonymous, The C. elegans Sequencing Consortium.
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: C89046
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <STO>
A:Cross-references: GB:chr_V; PIDN:AAB09171.1; PID:g1572829; GSPDB:GN00023; CESP:C10G8.4
C:Genetics:
A:Gene: C10G8.4
A:Map position: 5

Query Match 19.4%; Score 94.5; DB 2; Length 98;
Best Local Similarity 28.8%; Pred. No. 0.039;
Matches 23; Conservative 13; Mismatches 17; Indels 27; Gaps 6;

Qy 5 QCGENEKYDSCGSKCEDKKCKYDGVVEEDDEPN-----VPLVVRVCHODCVCEGFYRN 59
Db 39 RCPSENEEFKSCGT-ACEPTC-----QNPNPQVCTIQCLNVCO-----CSQGFVRG 83

Qy 60 KDKCVSAEDCELDNMDFIY 79
Db 84 PNG-CVPPQDC-----FVY 96

RESULT 14
T42737
gp330 protein precursor - rat
N:Alternate names: megalin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
C:Accession: T42737
R:Saito, A.; Pietromonaco, S.; Loo, A.K.C.; Farquhar, M.G.
Proc. Natl. Acad. Sci. U.S.A. 91, 9725-9729, 1994
A:Title: Complete cloning and sequencing of rat gp330/megalin, a distinctive member of ti
A:Reference number: A58173; MUID:95024033; PMID:7937880
A:Accession: T42737
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-4660 <SAI>
A:Cross-references: EMBL:L34049; NID:g561852; PID:g561853; PIDN:AAA51369.1
A:Experimental source: Strain Sprague-Dawley; kidney
C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-4660/Product: gp330 protein #status predicted <MAT>

Query Match 19.4%; Score 94.5; DB 2; Length 4660;
Best Local Similarity 30.0%; Pred. No. 0.97;
Matches 30; Conservative 11; Mismatches 24; Indels 35; Gaps 8;

Qy 6 CGENEKYDSC-----GSKEC---DKKCKYDGVVEE---EDDEPNV-----PCLVRY--- 45
Db 256 CESNQSHRRCPREWACPGSGRCISIDKVC--DGVDPCEBDDNNVTSGRTCGMGCVSV 313

Qy 46 -----CHQ-----DCVCEGFYRNKDDK---CVSAEDCEL 72
Db 314 LNCEYQCHQTFPGEGCFPCPGHIINSNDRSTCTDFDDCQI 353

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 / Search time 5.45455 Seconds
(without alignments)
116.399 Million cell updates/sec

Title: US-09-498-556C-85
Perfect score: 4
Sequence: 1 XXXX 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	1	20	AA1980.DAT
2	0	0.0	1	22	AA1981.DAT
3	0	0.0	1	22	AA1982.DAT
4	0	0.0	1	22	AA1983.DAT
5	0	0.0	1	22	AA1984.DAT
6	0	0.0	1	22	AA1985.DAT
7	0	0.0	1	22	AA1986.DAT
8	0	0.0	1	22	AA1987.DAT
9	0	0.0	1	22	AA1988.DAT

10	0	0.0	1	22	AA198447	Human peptide #172
11	0	0.0	1	22	AA198447	Human nonconservat
12	0	0.0	1	22	AA198447	Human nonconservat
13	0	0.0	1	22	AA198447	Human nonconservat
14	0	0.0	1	22	AA198447	Human nonconservat
15	0	0.0	1	22	AA198447	Human nonconservat
16	0	0.0	1	22	AA198447	Human nonconservat
17	0	0.0	1	22	AA198447	Human nonconservat
18	0	0.0	1	22	AA198447	Human nonconservat
19	0	0.0	1	22	AA198447	Human nonconservat
20	0	0.0	1	22	AA198447	Human nonconservat
21	0	0.0	1	22	AA198447	Human nonconservat
22	0	0.0	1	22	AA198447	Human nonconservat
23	0	0.0	1	22	AA198447	Human nonconservat
24	0	0.0	1	22	AA198447	Human nonconservat
25	0	0.0	1	22	AA198447	Human nonconservat
26	0	0.0	1	22	AA198447	Human nonconservat
27	0	0.0	1	22	AA198447	Human nonconservat
28	0	0.0	1	22	AA198447	Human nonconservat
29	0	0.0	1	22	AA198447	Human nonconservat
30	0	0.0	1	22	AA198447	Human nonconservat
31	0	0.0	1	22	AA198447	Human nonconservat
32	0	0.0	1	22	AA198447	Human nonconservat
33	0	0.0	1	22	AA198447	Human nonconservat
34	0	0.0	1	22	AA198447	Human nonconservat
35	0	0.0	1	22	AA198447	Human nonconservat
36	0	0.0	1	22	AA198447	Human nonconservat
37	0	0.0	1	22	AA198447	Human nonconservat
38	0	0.0	1	22	AA198447	Human nonconservat
39	0	0.0	1	22	AA198447	Human nonconservat
40	0	0.0	1	22	AA198447	Human nonconservat
41	0	0.0	1	22	AA198447	Human nonconservat
42	0	0.0	1	22	AA198447	Human nonconservat
43	0	0.0	1	22	AA198447	Human nonconservat
44	0	0.0	1	22	AA198447	Human nonconservat
45	0	0.0	1	22	AA198447	Human nonconservat

ALIGNMENTS

RESULT 1
AA198447
ID AA198447 standard; peptide; 1 AA.
XX AA198447
AC AA198447
XX AA198447
DT 01-DEC-1999 (first entry)
XX Immunogenic peptide having a human leukocyte antigen binding motif #1263.
DE Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
KW Immune response; T cell activation; major histocompatibility complex;
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
KW vaccine; immunisation.
XX Synthetic.
OS Homo sapiens.
XX WO9945954-A1.
XX 16-SEP-1999.
XX 13-MAR-1998; 98WO-US05039.
XX 13-MAR-1998; 98WO-US05039.
XX (EPIM-) EPIMUNE INC.
XX Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX WPI; 1999-551214/46.

CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 H 1

RESULT 4
 ABB66810
 ID ABB66810 standard; Protein; 1 AA.

XX ABB66810;

AC ABB66810;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 27222.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

PN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL10913.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 genes from Drosophila and for elucidating cell signalling and cell-cell
 interactions -

XX Disclosure; SEQ ID NO 27222; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 H 1

RESULT 5
 ABG02941
 ID ABG02941 standard; Protein; 1 AA.

XX ABG02941;

DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #2932.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS67128.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

XX Claim 20; SEQ ID NO 33300; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

```

Db      1 M 1

RESULT 6
AAM97643
ID AAM97643 standard; Peptide; 1 AA.
XX
AC AAM97643;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human peptide #918 encoded by a SNP oligonucleotide.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Disclosure; Page 3869; 4143pp; English.
XX
CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythromatosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms.
XX
SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 C 1

RESULT 7
AAM97834
ID AAM97834 standard; Peptide; 1 AA.
XX
AC AAM97834;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human peptide #1109 encoded by a SNP oligonucleotide.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Disclosure; Page 3911; 4143pp; English.
XX
CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythromatosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms.
XX
SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 E 1

RESULT 8
AAM97974

```


ID AM97974 standard; Peptide; 1 AA.
 XX
 AC
 XX
 DT
 XX
 DT
 XX
 DE
 XX
 DE
 Human peptide #1249 encoded by a SNP oligonucleotide.
 XX
 XX
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; cancer;
 KW amyloid protein; angiopeptin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesis; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200147944-A2.
 XX
 PD 05-JUL-2001.
 XX
 PF 28-DEC-2000; 2000WO-US35498.
 XX
 PR 28-DEC-1999; 99US-0173419.
 XX
 PR 27-DEC-2000; 2000US-0173419.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI; 2001-465210/50.
 XX
 PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections -
 XX
 PS Disclosure; Page 3941; 4143pp; English.
 XX
 CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
 CC encoding polymorphic variants of proteins related to amylases, amyloid
 CC proteins, angiopeptin, apoptosis related proteins, cadherin, cyclin,
 CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
 CC complement related proteins, cytochromes, kinesis, cytokines,
 CC interferons, interleukins, G-protein coupled receptors and thioesterases.
 CC The present sequence is a peptide encoded by one such oligonucleotide.
 CC The oligonucleotides and the peptides encoded by them may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate expression of the proteins listed above. Disorders that may
 CC be prevented, diagnosed and/or treated include multifactorial diseases
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
 CC system and an infection of pathogenic organisms.
 XX
 SQ Sequence 1 AA;
 Query Match 0.0%; Score 0; DB 22; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 X 1
 DB 1 L 1
 RESULT 9
 AM98354
 ID AM98354 standard; Peptide; 1 AA.
 XX
 AC
 XX
 DT 24-JAN-2002 (first entry)
 XX
 DE
 XX
 DE Human peptide #1629 encoded by a SNP oligonucleotide.
 XX
 XX
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; cancer;
 KW amyloid protein; angiopeptin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesis; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200147944-A2.
 XX
 PD 05-JUL-2001.
 XX
 PF 28-DEC-2000; 2000WO-US35498.
 XX
 PR 28-DEC-1999; 99US-0173419.
 XX
 PR 27-DEC-2000; 2000US-0173419.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI; 2001-465210/50.
 XX
 PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections -
 XX
 PS Disclosure; Page 4025; 4143pp; English.
 XX
 CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
 CC encoding polymorphic variants of proteins related to amylases, amyloid
 CC proteins, angiopeptin, apoptosis related proteins, cadherin, cyclin,
 CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
 CC complement related proteins, cytochromes, kinesis, cytokines,
 CC interferons, interleukins, G-protein coupled receptors and thioesterases.
 CC The present sequence is a peptide encoded by one such oligonucleotide.
 CC The oligonucleotides and the peptides encoded by them may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate expression of the proteins listed above. Disorders that may
 CC be prevented, diagnosed and/or treated include multifactorial diseases
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
 CC system and an infection of pathogenic organisms.
 XX
 SQ Sequence 1 AA;
 Query Match 0.0%; Score 0; DB 22; Length 1;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 X 1
 DB 1 L 1
 RESULT 10
 AM98447
 ID AM98447 standard; Peptide; 1 AA.
 XX
 AC
 XX
 DT 24-JAN-2002 (first entry)
 XX

DE Human peptide #1722 encoded by a SNP oligonucleotide.

KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;

KW neuroprotective; antimicrobial; gene therapy; vaccine; cancer;

KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;

KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;

KW complement related protein; cytochrome; kinesin; cytokine; interferon;

KW interleukin; G-protein coupled receptor; thioesterase; inflammation;

KW multifactorial disease; autoimmune disease; infection;

KW nervous system disease.

XX Homo sapiens.

OS WO200147944-A2.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-US35498.

XX 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2001-465210/50.

DR Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,

XX oncogenes and histones, useful for diagnosing and treating, e.g.

PT cancer, autoimmune diseases and infections

PT Disclosure; Page 4045; 4143pp; English.

XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)

CC encoding polymorphic variants of proteins related to amyloid

CC proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin,

CC polymerase, oncogenes, histones, kinases, colony stimulating factors,

CC complement related proteins, cytochromes, kinesins, cytokines,

CC interferons, interleukins, G-protein coupled receptors and thioesterases.

CC The present sequence is a peptide encoded by one such oligonucleotide.

CC The oligonucleotides and the peptides encoded by them may be used in the

CC prevention, diagnosis and treatment of diseases associated with

CC inappropriate expression of the proteins listed above. Disorders that may

CC be prevented, diagnosed and/or treated include multifactorial diseases

CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid

CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus

CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,

CC brain, breast, colon and kidney, leukaemia), diseases of the nervous

CC system and an infection of pathogenic organisms.

XX

SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;

Best Local Similarity 0.0%; Pred.No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 L 1

RESULT 11

AAM53218

ID AAM53218 standard; Peptide; 1 AA.

XX AAM53218;

AC AAM53218;

XX 09-NOV-2001 (first entry)

DT Human nonconservative amino acid changing SNP related peptide SEQ:6913.

DE Human; single nucleotide polymorphism; SNP; genome; gene therapy;

XX

KW

protein therapy; vaccine; probe; diagnostic assay; detection;

quantitation; restorative therapy; polymorphic.

Homo sapiens.

WO200140521-A2.

07-JUN-2001.

30-NOV-2000; 2000WO-US32758.

30-NOV-1999; 99US-0168138.

29-NOV-2000; 2000US-0726173.

(CURA-) CURAGEN CORP.

Shimkets RA, Leach M;

WPI; 2001-356160/37.

Polymorphic nucleic acid sequences, useful in genetic testing and

therapy

Claim 29; Page 2619; 2653pp; English.

AAL73060 to AAL79867 represent isolated human polymorphic polynucleotide

sequences (I), which contain single nucleotide polymorphisms (SNPs).

AAM53114 to AAM53329 represent peptides related to human polymorphic

polynucleotide sequences. The sequences can be used in gene and protein

therapy, and in vaccine production. (I) and the polypeptides encoded by

them may be used in the prevention, diagnosis and treatment of diseases

associated with inappropriate expression of polymorphic polypeptides.

For example, (I) may be used to treat disorders by rectifying mutations

or deletions in a patient's genome that affect the activity of

polypeptides by expressing inactive proteins or to supplement the

patients own production of polypeptide. Additionally, (I) and its

complementary sequences may also be used as DNA probes in diagnostic

assays to detect and quantitate the presence of similar nucleic acids

in samples, and therefore which patients may be in need of restorative

therapy. The polypeptides encoded by (I) may be used as antigens in the

production of antibodies specific for polymorphic polypeptides. The

antibodies may also be used to down regulate expression and activity.

The antibodies may also be used as diagnostic agents for detecting the

presence of polymorphic polypeptides in samples.

XX Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 H 1

RESULT 12

AAM53219

ID AAM53219 standard; Peptide; 1 AA.

XX AAM53219;

AC AAM53219;

XX 09-NOV-2001 (first entry)

DT Human nonconservative amino acid changing SNP related peptide SEQ:6914.

DE Human; single nucleotide polymorphism; SNP; genome; gene therapy;

KW protein therapy; vaccine; probe; diagnostic assay; detection;

KW quantitation; restorative therapy; polymorphic.

XX Homo sapiens.

OS WO200140521-A2.

XX

XX PD 07-JUN-2001.
XX PF 30-NOV-2000; 2000WO-US32758.
XX PR 30-NOV-1999; 99US-0168138.
XX PR 29-NOV-2000; 2000US-0726173.
XX PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach M;
XX PI WPI; 2001-356160/37.
XX PT Polymorphic nucleic acid sequences, useful in genetic testing and
XX PT therapy -
XX PS Claim 29; Page 2619; 2653pp; English.
XX CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
XX CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
XX CC AAM53114 to AAM53329 represent peptides related to human polymorphic
XX CC polynucleotide sequences. The sequences can be used in gene and protein
XX CC therapy, and in vaccine production. (I) and the polypeptides encoded by
XX CC them may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate expression of polymorphic polypeptides.
XX CC For example, (I) may be used to treat disorders by rectifying mutations
XX CC or deletions in a patient's genome that affect the activity of
XX CC polypeptides by expressing inactive proteins or to supplement the
XX CC patients own production of polypeptide. Additionally, (I) and its
XX CC complementary sequences may also be used as DNA probes in diagnostic
XX CC assays to detect and quantitate the presence of similar nucleic acids
XX CC in samples, and therefore which patients may be in need of restorative
XX CC therapy. The polypeptides encoded by (I) may be used as antigens in the
XX CC production of antibodies specific for polymorphic polypeptides. The
XX CC antibodies may also be used to down regulate expression and activity.
XX CC The antibodies may also be used as diagnostic agents for detecting the
XX CC presence of polymorphic polypeptides in samples.
XX SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 H 1
RESULT 13
AAM53290
ID AAM53290 standard; Peptide; 1 AA.
XX AC AAM53290;
XX DT 09-NOV-2001 (first entry)
XX DE Human nonconservative amino acid changing SNP related peptide SEQ:6985.
XX KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
XX KW protein therapy; vaccine; probe; diagnostic assay; detection;
XX KW quantitation; restorative therapy; polymorphic.
XX OS Homo sapiens.
XX PN WO200140521-A2.
XX PD 07-JUN-2001.
XX PF 30-NOV-2000; 2000WO-US32758.
XX PR 30-NOV-1999; 99US-0168138.
XX PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach M;
XX PI WPI; 2001-356160/37.

PR 29-NOV-2000; 2000US-0726173.
XX (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach M;
XX XX WPI; 2001-356160/37.
XX PT Polymorphic nucleic acid sequences, useful in genetic testing and
XX PT therapy -
XX PS Claim 29; Page 2641; 2653pp; English.
XX CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
XX CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
XX CC AAM53114 to AAM53329 represent peptides related to human polymorphic
XX CC polynucleotide sequences. The sequences can be used in gene and protein
XX CC therapy, and in vaccine production. (I) and the polypeptides encoded by
XX CC them may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate expression of polymorphic polypeptides.
XX CC For example, (I) may be used to treat disorders by rectifying mutations
XX CC or deletions in a patient's genome that affect the activity of
XX CC polypeptides by expressing inactive proteins or to supplement the
XX CC patients own production of polypeptide. Additionally, (I) and its
XX CC complementary sequences may also be used as DNA probes in diagnostic
XX CC assays to detect and quantitate the presence of similar nucleic acids
XX CC in samples, and therefore which patients may be in need of restorative
XX CC therapy. The polypeptides encoded by (I) may be used as antigens in the
XX CC production of antibodies specific for polymorphic polypeptides. The
XX CC antibodies may also be used to down regulate expression and activity.
XX CC The antibodies may also be used as diagnostic agents for detecting the
XX CC presence of polymorphic polypeptides in samples.
XX SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 L 1
RESULT 14
AAM53291
ID AAM53291 standard; Peptide; 1 AA.
XX AC AAM53291;
XX DT 09-NOV-2001 (first entry)
XX DE Human nonconservative amino acid changing SNP related peptide SEQ:6986.
XX KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
XX KW protein therapy; vaccine; probe; diagnostic assay; detection;
XX KW quantitation; restorative therapy; polymorphic.
XX OS Homo sapiens.
XX PN WO200140521-A2.
XX PD 07-JUN-2001.
XX PF 30-NOV-2000; 2000WO-US32758.
XX PR 30-NOV-1999; 99US-0168138.
XX PR 29-NOV-2000; 2000US-0726173.
XX PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach M;
XX PI WPI; 2001-356160/37.

DR WPI; 2001-356160/37.
XX Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
XX
PS Claim 29; Page 2641; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 X 1
Db 1 L 1
Search completed: December 8, 2003, 09:16:39
Job time : 5.45455 secs

DR WPI; 2001-356160/37.
XX Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
XX
PS Claim 29; Page 2641; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 1 AA;
Query Match 0.0%; Score 0; DB 22; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 X 1
Db 1 L 1

RESULT 15
AAM53328
ID AAM53328 standard; Peptide; 1 AA.
XX
AC AAM53328;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human nonconservative amino acid changing SNP related peptide SEQ.7023.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP..
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
XX Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
XX Claim 29; Page 2652; 2653pp; English.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 4.20202 Seconds
(without alignments)
245.646 Million cell updates/sec

Title: US-09-498-556C-85
Perfect score: 4
Sequence: 1 XXXX 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	4	11 Q08433	Q08433 rattus sp.
2	0	0.0	5	2 P83073	P83073 bacillus ce
3	0	0.0	5	10 Q99007	Q99007 hordeum vul
4	0	0.0	5	13 P83308	P83308 gallus gall
5	0	0.0	6	10 P82181	P82181 spinacia ol
6	0	0.0	6	10 P82541	P82541 spinacia ol
7	0	0.0	6	10 P82182	P82182 spinacia ol
8	0	0.0	7	2 Q8KMS3	Q8KMS3 klebsiella
9	0	0.0	7	2 Q47505	Q47505 escherichia
10	0	0.0	7	2 P70804	P70804 azotobacter
11	0	0.0	7	2 O50556	O50556 actinobacil
12	0	0.0	7	2 Q47477	Q47477 escherichia
13	0	0.0	7	2 Q47029	Q47029 enterobacte
14	0	0.0	7	2 P72081	P72081 nocardia la
15	0	0.0	7	2 Q54248	Q54248 streptomyce
16	0	0.0	7	2 Q8KMS9	Q8KMS9 enterobacte

17	0	0.0	7	2	Q34028	O34028 sphingomona
18	0	0.0	7	2	O07354	O07354 synechococc
19	0	0.0	7	2	O8GL12	O8GL12 borrelia bu
20	0	0.0	7	2	O8GL04	O8GL04 borrelia bu
21	0	0.0	7	2	O8GL00	O8GL00 borrelia bu
22	0	0.0	7	4	Q15903	Q15903 homo sapien
23	0	0.0	7	4	Q8NH7	Q8NH7 homo sapien
24	0	0.0	7	4	Q15897	Q15897 homo sapien
25	0	0.0	7	6	Q28742	Q28742 oryctolagus
26	0	0.0	7	8	P92214	P92214 amblyopyrum
27	0	0.0	7	8	P92393	P92393 hordeum vul
28	0	0.0	7	8	P92403	P92403 lophopyrum
29	0	0.0	7	8	P92427	P92427 peridictyon
30	0	0.0	7	8	O99182	O99182 gnatholebia
31	0	0.0	7	8	P92430	P92430 aegilops ta
32	0	0.0	7	8	P92221	P92221 bromus iner
33	0	0.0	7	8	O98866	O98866 spinacia ol
34	0	0.0	7	8	P92425	P92425 pseudoroegn
35	0	0.0	7	8	P92381	P92381 hordeum bra
36	0	0.0	7	8	P92387	P92387 henrardia p
37	0	0.0	7	8	P92210	P92210 agropyron c
38	0	0.0	7	8	P92440	P92440 thinopyrum
39	0	0.0	7	8	P92218	P92218 australopyr
40	0	0.0	7	8	P92390	P92390 heteranthel
41	0	0.0	7	8	P92372	P92372 haynaldia v
42	0	0.0	7	8	P92442	P92442 taeniatheru
43	0	0.0	7	8	P92226	P92226 crithopsis
44	0	0.0	7	8	Q8MFY6	Q8MFY6 taraxacum (
45	0	0.0	7	8	Q95945	Q95945 saccharomyc

ALIGNMENTS

RESULT 1

Q08433 PRELIMINARY; PRT; 4 AA.
AC Q08433;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Bilirubin UDP-glucuronosyltransferase (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Gunn;
RX MEDLINE=91282758; PubMed=1840486;
RA Sato H., Aono S., Kashiwamata S., Koiwai O.;
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
hyperbilirubinemic Gunn rat."
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
DR EMBL; S38636; AAB19259.1; -;
KW Transferase.
FT NON_TER 1
SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 0.0%; Score 0; DB 11; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 N 1

RESULT 2

P83073 PRELIMINARY; PRT; 5 AA.
ID P83073
AC P83073;
DT 01-OCT-2001 (TREMBLrel. 18, Created)

DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE 88 kDa protein (Fragment).
 OS Bacillus cereus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1396;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=NCIMB 11796;
 RA Browne N., Dowds B.C.A.;
 RL Submitted (JUL-2001) to the SWISS-PROT data bank.
 FT NON TER 5
 SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F00000 CRC64;
 Query Match 0.0%; Score 0; DB 2; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 X 1
 Db 4 T 4
 RESULT 3
 Q99007
 ID Q99007 PRELIMINARY; PRT; 5 AA.
 AC Q99007;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Alpha amylase (Fragment).
 GN AMY1 GENE
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91329704; PubMed=1831055;
 RA Jacobsen J.V., Close T.J.;
 RT "Control of transient expression of chimaeric genes by gibberellic
 RT acid and abscisic acid in protoplasts prepared from mature barley
 RT aleurone layers";
 RL Plant Mol. Biol. 16:713-721(1991).
 DR EMBL; X54643; CAA38455.1; -;
 FT NON TER 5
 SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRC64;
 Query Match 0.0%; Score 0; DB 10; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 X 1
 Db 2 A 2
 RESULT 4
 P83308
 ID P83308 PRELIMINARY; PRT; 5 AA.
 AC P83308;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE FMRamide-like neuropeptide (LPLRF-amide).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.

RC TISSUE=Brain;
 RX PubMed=6137771;
 RA Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;
 RT "A novel active pentapeptide from chicken brain identified by
 RT antibodies to FMRamide";
 RL Nature 305:328-330(1983).
 CC -1- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 CC Neuropeptide.
 KW Neuropeptide.
 SQ SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64;
 Query Match 0.0%; Score 0; DB 13; Length 5;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 X 1
 Db 1 L 1
 RESULT 5
 P82181
 ID P82181 PRELIMINARY; PRT; 6 AA.
 AC P82181;
 DT 01-JUN-2000 (TrEMBLrel. 14, Created)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Chloroplast 50S ribosomal protein L10 beta (Fragment).
 OS Spinacia oleracea (Spinach).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=cv. ALVARO; TISSUE=Leaf;
 RX MEDLINE=20435798; PubMed=10874046;
 RA Yamaguchi K., Subramanian A.R.;
 RT "The plastid ribosomal proteins. Identification of all the proteins in
 RT the 50 S subunit of an organelle ribosome (chloroplast).";
 RL J. Biol. Chem. 275:28466-28482(2000).
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
 DR InterPro; IPR001790; Ribosomal_L10.
 DR Pfam; PF00466; Ribosomal_L10; PARTIAL.
 DR PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
 KW Ribosomal protein; Chloroplast; rRNA-binding.
 FT NON TER 6
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;
 Query Match 0.0%; Score 0; DB 10; Length 6;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 X 1
 Db 1 A 1
 RESULT 6
 P82541
 ID P82541 PRELIMINARY; PRT; 6 AA.
 AC P82541;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Chloroplast 30S ribosomal protein S19 beta (Fragment).
 OS Spinacia oleracea (Spinach).

OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
 RC STRAIN=cv. ALVARO; TISSUE=Leaf;
 RA MEDLINE=20435797; PubMed=10874039;
 RX Yamaguchi K., von Knoblauch K., Subramanian A.R.;
 RT "The plastid ribosomal proteins. Identification of all the proteins in
 RL the small subunit of an organelle ribosome (chloroplast).";
 J. Biol. Chem. 275:28466-28482(2000).
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC -1- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.
 CC -1- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.
 CC -1- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA
 CC FORM IS THE MINOR BASIC FORM.
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.
 CC -1- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
 DR InterPro: IPR002222; Ribosomal S19.
 DR Pfam: PF00203; Ribosomal S19; PARTIAL.
 DR PRINTS; PR00975; RIBOSOMALS19; PARTIAL.
 DR PROSITE; PS00323; RIBOSOMAL S19; PARTIAL.
 KW Ribosomal protein; Chloroplast; rRNA-binding.
 FT NON TER 6 6
 SQ SEQUENCE 6 AA; 732 MW; 63333735A411C000 CRC64;
 Query Match 0.0%; Score 0; DB 10; Length 6;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 DB 1 T 1
 RESULT 7
 ID P82182 PRELIMINARY; PRT; 6 AA.
 AC P82182;
 DT 01-JUN-2000 (TrEMBLrel. 14, Created)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Chloroplast 50S ribosomal protein L10 gamma (fragment).
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=cv. ALVARO; TISSUE=Leaf;
 RX MEDLINE=20435798; PubMed=10874046;
 RA Yamaguchi K., Subramanian A.R.;
 RT "The plastid ribosomal proteins. Identification of all the proteins in
 RL the 50 S subunit of an organelle ribosome (chloroplast).";
 J. Biol. Chem. 275:28466-28482(2000).
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
 DR InterPro: IPR001790; Ribosomal L10.
 DR Pfam: PF00466; Ribosomal L10; PARTIAL.
 DR PROSITE; PS01109; RIBOSOMAL L10; PARTIAL.
 KW Ribosomal protein; Chloroplast; rRNA-binding.
 FT NON TER 6 6
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 6;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 DB 1 A 1
 RESULT 8
 ID Q8KWS3 PRELIMINARY; PRT; 7 AA.
 AC Q8KWS3;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Putative Merr2 protein.
 GN MERR2.
 OS Klebsiella sp. LS13-39.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Klebsiella.
 OX NCBI_TaxID=143776;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LS13-39;
 RX MEDLINE=21604134; PubMed=11763242;
 RA Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,
 RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,
 RA Yurieva O.V., Nikiforov V.G.;
 RT "Mercury resistance transposons of Gram-negative environmental
 RT bacteria and their classification.";
 RL Res. Microbiol. 152:811-822(2001).
 DR EMBL; AJ302776; CAC82975.1; -
 SQ SEQUENCE 7 AA; 608 MW; 6DC1B5BDD87DD6F0 CRC64;
 Query Match 0.0%; Score 0; DB 2; Length 7;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 DB 2 A 2
 RESULT 9
 ID Q47505 PRELIMINARY; PRT; 7 AA.
 AC Q47505;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE MccA protein.
 GN MCCA.
 OS Escherichia coli.
 OG Plasmid pMccC7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96099297; PubMed=8522520;
 RA Gonzalez-Pastor J.E., San Millan J.L., Castilla M.A., Moreno F.;
 RT "Structure and organization of plasmid genes required to produce the
 RL translation inhibitor microcin C7.";
 J. Bacteriol. 177:7131-7140(1995).
 DR EMBL; X57583; CAA40808.1; -
 KW Plasmid.
 SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;
 Query Match 0.0%; Score 0; DB 2; Length 7;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 3 T 3

RESULT 10

P70804 PRELIMINARY; PRT; 7 AA.
AC P70804; 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE AlgT protein (Fragment).
GN ALGT.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=; PubMed=8830682;
RX MEDLINE=96427318; Ertesvag H., Valla S.;
RA Rehm B.H.A., "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is
RT part of an alg gene cluster physically organized in a manner similar
RT to that in Pseudomonas aeruginosa.";
RL J. Bacteriol. 178:5884-5889(1996).
DR EMBL; X87973; CAA61230.1; -.
FT NON TER 1
SQ SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 2 T 2

RESULT 11

O50556 PRELIMINARY; PRT; 7 AA.
AC O50556; 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE GLYA (Fragment).
GN GLYA.
OS Actinobacillus actinomycetemcomitans (Haemophilus
OS actinomycetemcomitans).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33384;
RX MEDLINE=96355846; PubMed=8751884;
RA Kolodrubetz D., Spitznagel J. Jr., Wang B., Phillips L.H., Jacobs C.,
RA Kraig E.;
RT "cis Elements and trans factors are both important in strain-specific
RT regulation of the leukotoxin gene in Actinobacillus
RT actinomycetemcomitans.";
RL Infect. Immun. 64:3451-3460(1996).
DR EMBL; U51862; AAB88721.1; -.
FT NON TER 1
SQ SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 7 A 7

RESULT 12

Q47477 PRELIMINARY; PRT; 7 AA.
AC Q47477; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Tpi protein (Fragment).
GN TPI.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE OF 7-7 FROM N.A.
RX MEDLINE=85203917; PubMed=3158524;
RA Hellinga H.W., Evans P.R.;
RT "Nucleotide sequence and high-level expression of the major
RT Escherichia coli phosphofructokinase.";
RL Eur. J. Biochem. 149:363-373(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Evans P.;
RL Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.
DR EMBL; X02519; CAA26359.1; -.
FT NON TER 1
SQ SEQUENCE 7 AA; 773 MW; 7416D33DDDBD1DB0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 A 1

RESULT 13

Q47029 PRELIMINARY; PRT; 7 AA.
AC Q47029; 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Aad A1 protein (Fragment).
GN Aad A1.
OS Enterobacter cloacae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Enterobacter.
OX NCBI_TaxID=550;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94079349; PubMed=8257126;
RA Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;
RT "Analysis of the sac(3)-Via gene encoding a novel 3-N-
RT acetyltransferase.";
RL Antimicrob. Agents Chemother. 37:2074-2079(1993).
DR EMBL; M88012; AAA16193.1; -.
FT NON TER 1
SQ SEQUENCE 7 AA; 744 MW; 633862D2C321A030 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 2 T 2

RESULT 14

P72081
 ID P72081 PRELIMINARY; PRT; 7 AA.
 AC P72081;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE 3'-methylcephem hydroxylase (Fragment).
 GN CEFF.
 OS Nocardia lactamdurans.
 OC Bacteria; Actinobacteridae; Actinobacteriales; Actinomycetales;
 OC Pseudonocardineae; Pseudonocardaceae; Amycolatopsis.
 OX NCBI_TaxID=1913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96009872; PubMed=7557411;
 RA Coque J., Perez-Llarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,
 RA Liras P.;
 RT "Characterization of the cmh genes of Nocardia lactamdurans and
 RT Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
 RT O-carbamoyltransferase for cephamycin biosynthesis.";
 RL Gene 162:21-27(1995).
 DR EMBL; Z21682; CAA79797.1; --
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 746 MW; 75A1B2CDD1E681C0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
 Best Local Similarity 0.0%; Pred.No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
 Db 1 T 1

RESULT 15

Q54248
 ID Q54248 PRELIMINARY; PRT; 7 AA.
 AC Q54248;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE RplO protein (Fragment).
 GN RPL0.
 OS Streptomyces griseus.
 OC Bacteria; Actinobacteridae; Actinobacteriales; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1911;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=N2-3-11;
 RX MEDLINE=20011291; PubMed=10542330;
 RA Poehling S., Fiepersberg W., Wehmeier U.F.;
 RT "Analysis and regulation of the sec Y gene from Streptomyces griseus
 RT N2-3-11 and interaction of the SecY protein with the SecA protein.";
 RL Biochim. Biophys. Acta 1447:298-302(1999).
 DR EMBL; X95915; CAA65160.1; --
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 760 MW; 72C72B01B2D1B2A0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;
 Best Local Similarity 0.0%; Pred.No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
 Db 2 T 2

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 1.05051 Seconds
(without alignments)
179.063 Million cell updates/sec

Title: US-09-498-556C-85
Perfect score: 4
Sequence: 1 XXXX 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	3	GRWM_HUMAN	P01157 homo sapien
2	0	0.0	3	LUXE_VIBFI	P24272 vibrio fisc
3	0	0.0	3	THYL_PIG	P01151 sus scrofa
4	0	0.0	4	ACH1_ACHFU	P35904 achatina fu
5	0	0.0	4	DCML_PSECH	P19916 pseudomonas
6	0	0.0	4	DCMS_PSECH	P19918 pseudomonas
7	0	0.0	4	EOSI_HUMAN	P02731 homo sapien
8	0	0.0	4	FAR3_HIRME	P42562 hirudo medi
9	0	0.0	4	FAR4_HIRME	P42563 hirudo medi
10	0	0.0	4	FFKA_ANTEL	P58705 anthopleura
11	0	0.0	4	FLRF_HIRME	P42561 hirudo medi
12	0	0.0	4	FLRN_ANTEL	P58707 anthopleura
13	0	0.0	4	FMRP_WACNI	P01162 macrocallis
14	0	0.0	4	FYRI_ANTEL	P58706 anthopleura
15	0	0.0	4	OCPI_OCTMI	P58648 octopus min
16	0	0.0	4	OCPI_OCTMI	P58649 octopus min
17	0	0.0	4	RM01_YEAST	P36515 saccharomyc
18	0	0.0	4	TUFT_HUMAN	P01858 homo sapien
19	0	0.0	5	AL14_CARMA	P81817 carcinus ma
20	0	0.0	5	BIOA_CITFR	P13071 citrobacter
21	0	0.0	5	BIOB_CITFR	P12997 citrobacter
22	0	0.0	5	BP71_BOTIN	P30425 bothrops in
23	0	0.0	5	E103_LITRU	P82099 litoria rub
24	0	0.0	5	E104_LITRU	P82100 litoria rub
25	0	0.0	5	FARP_ARTTR	P41853 artiposthi
26	0	0.0	5	PAP2_FARMA	P81864 pardachirus
27	0	0.0	5	PRCT_PERAM	P01373 periplaneta
28	0	0.0	5	PSK_DAUCA	P58261 daucus caro
29	0	0.0	5	RE11_LITRU	P82070 litoria rub
30	0	0.0	5	RE21_LITRU	P82071 litoria rub
31	0	0.0	5	RE31_LITRU	P82072 litoria rub
32	0	0.0	5	RE32_LITRU	P82073 litoria rub
33	0	0.0	5	SUGA_ACHDO	P19991 acheta dome

ALIGNMENTS

RESULT 1

GRWM_HUMAN STANDARD; PRT; 3 AA.
AC P01157;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Growth-modulating peptide.
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77162369; PubMed=858356;
RA Schlesinger D.H., Pickart L., Thaler M.M.;
RT "Growth-modulating serum tripeptide is glycyl-histidyl-lysine.";
RL Experientia 33:324-325(1977).
CC -!- MISCELLANEOUS: THIS SERUM TRIPEPTIDE HAS BEEN FOUND TO STIMULATE
CC GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.
SQ GO: GO:0001558; P:regulation of cell growth; NAS.
DR SEQUENCE 3 AA; 340 MW; 6331E81000000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 G 1

RESULT 2

LUXE_VIBFI STANDARD; PRT; 3 AA.
ID LUXE_VIBFI
AC P24272;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Long-chain-fatty-acid-luciferin-component ligase (EC 6.2.1.19) (Acyl-protein synthetase) (Fragment).
GN LUXE.
OS Vibrio fischeri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91072226; PubMed=2254256;
RA Swartzman E., Kapoor S., Graham A.F., Meighen B.A.;
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination site for the lux operon.";
RL J. Bacteriol. 172:6797-6802(1990).
CC -!- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.
CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS

CC SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.
 CC -1- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +
 CC an acyl-protein thioester.
 CC -1- PATHWAY: Bioluminescent fatty acid reduction system; second step.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M62812; ; NOT_ANNOTATED_CDS.
 CC Luminance; Ligase. 1
 CC NON_TER 1
 CC SEQUENCE 3 AA; 374 MW; 6AA3303000000000000 CRC64;
 CC
 CC Query Match 0.0%; Score 0; DB 1; Length 3;
 CC Best Local Similarity 0.0%; Pred. No. 0;
 CC Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC Qy 1 X 1
 CC Db 1 I 1
 CC
 CC RESULT 3
 CC THYL_PIG
 CC ID THYL_PIG STANDARD; PRT; 3 AA.
 CC AC P01151;
 CC DT 21-JUL-1986 (Rel. 01, Created)
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Thyloliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).
 CC OS Sus scrofa (Pig).
 CC OS Ovis aries (Sheep).
 CC OS Bombina orientalis (Oriental fire-bellied toad), and
 CC OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 CC OX NCBI_TaxID=9823, 9940, 8346, 8316;
 CC [1]
 CC RP SEQUENCE.
 CC RC SPECIES=Pig; TISSUE=Hypothalamus;
 CC RX MEDLINE=70136150; PubMed=4984938;
 CC RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;
 CC RT "Structure of porcine thyrotropin releasing hormone."
 CC RL Biochemistry 9:1103-1106(1970).
 CC [2]
 CC RP SYNTHESIS.
 CC RC SPECIES=Pig;
 CC RX MEDLINE=70039904; PubMed=4982117;
 CC RA Boler J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;
 CC RT "The identity of chemical and hormonal properties of the thyrotropin
 CC releasing hormone and pyroglutamyL-histidyl-proline amide."
 CC RL Biochem. Biophys. Res. Commun. 37:705-710(1969).
 CC [3]
 CC RP SEQUENCE.
 CC RC SPECIES=Sheep; TISSUE=Hypothalamus;
 CC RA Desiderio D.M. Jr., Burgus R., Dunn T.F., Vale W., Guillemin R.,
 CC RA Ward D.N.;
 CC RT "The elucidation of the primary structure of the hypothalamic thyroid
 CC stimulating hormone releasing factor of ovine origin by means of mass
 CC spectrometry."
 CC RL Org. Mass Spectrom. 5:221-228(1971).
 CC [4]
 CC RP SYNTHESIS.
 CC RC SPECIES=Sheep;
 CC RX MEDLINE=70163386; PubMed=4985794;
 CC RA Burgus R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,
 CC RA Guillemin R.;
 CC RT "Characterization of ovine hypothalamic hypophysiotropic

RT TSH-releasing factor."
 RL Nature 226:321-325(1970).
 RN [5]
 RN SEQUENCE.
 RC SPECIES=B.orientalis; TISSUE=Skin;
 RX MEDLINE=76138399; PubMed=815011;
 RA Yasuhara T., Nakajima T.;
 RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin."
 RL Chem. Pharm. Bull. 23:3301-3303(1975).
 RN [6]
 RN SEQUENCE.
 RC SPECIES=N.viridescens;
 RX MEDLINE=75035605; PubMed=4214528;
 RA Grimm-Joergensen Y., McKelvy J.F.;
 RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus
 RT viridescens) brain in vitro. Isolation and characterization of
 RT thyrotropin releasing factor."
 RL J. Neurochem. 23:471-478(1974).
 CC -1- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF TSH
 CC IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/
 CC NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.
 CC
 CC DR PIR; A90919; RHTDIO.
 CC DR PIR; A92971; A92971.
 CC DR PIR; A93750; RHSHT.
 CC KW Amidation; Pyrrolidone carboxylic acid.
 CC FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 CC FT MOD_RES 3 3 AMIDATION.
 CC SQ SEQUENCE 3 AA; 380 MW; 7761F6B0000000000000 CRC64;
 CC
 CC Query Match 0.0%; Score 0; DB 1; Length 3;
 CC Best Local Similarity 0.0%; Pred. No. 0;
 CC Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC Qy 1 X 1
 CC Db 1 Q 1
 CC
 CC RESULT 4
 CC ACHI_ACHFU
 CC ID ACHI_ACHFU STANDARD; PRT; 4 AA.
 CC AC P35904;
 CC DT 01-JUN-1994 (Rel. 29, Created)
 CC DT 01-JUN-1994 (Rel. 29, Last sequence update)
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)
 CC DE Achatina-I.
 CC OS Achatina fulica (Giant African snail).
 CC OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 CC OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
 CC OX NCBI_TaxID=6530;
 CC [1]
 CC RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
 CC RC STRAIN=Ferussac; TISSUE=Ganglion;
 CC RX MEDLINE=89273551; PubMed=2597281;
 CC RA Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,
 CC RA Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,
 CC RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;
 CC RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina
 CC fulica Ferussac containing a D-amino acid residue."
 CC RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
 CC [2]
 CC RP CHARACTERIZATION.
 CC RC STRAIN=Ferussac; TISSUE=Heart atrium;
 CC RX MEDLINE=91264856; PubMed=1675568;
 CC RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,
 CC RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;
 CC RT "Purification of achatin-I from the atria of the African giant snail,
 CC Achatina fulica, and its possible function."
 CC RL Biochem. Biophys. Res. Commun. 177:847-853(1991).
 CC [3]
 CC RP X-RAY CRYSTALLOGRAPHY.
 CC RX MEDLINE=93014529; PubMed=1399265;
 CC RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,

RA Iwashita T., Nomoto K.;
 RT "Crystal structure and molecular conformation of achainin-I
 RT (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a
 RT D-amino acid residue.";
 RL Int. J. Pept. Protein Res. 39:258-264(1992).
 CC -1- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY
 CC AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY
 CC NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE
 CC HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.
 DR PIR; A32480; A32480.
 KW Hormone; D-amino acid.
 FT MOD_RES 2 D-PHENYLALANINE.
 SQ SEQUENCE 4 AA; 408 MW; 6AADD9C8100000000 CRC64;
 Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 1; Gaps 0;
 QY 1 X 1
 Db 3 A 3
 RESULT 5
 DCML_PSECH STANDARD; PRT; 4 AA.
 ID DCML_PSECH
 AC P19916;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
 DE dehydrogenase subunit L) (CO-DH L) (Fragment).
 GN CUTL.
 OS Pseudomonas carboxydohydrogena.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae.
 OX NCBI_TaxID=290;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=90055678; PubMed=2818128;
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
 RT "Homology and distribution of CO dehydrogenase structural genes in
 RT carboxydohydrogenic bacteria.";
 RL Arch. Microbiol. 152:335-341(1989).
 CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
 CC dioxide.
 CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
 CC acceptor.
 CC -1- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.
 CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
 CC SMALL.
 DR PIR; P10146; P10146.
 KW Oxidoreductase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S.
 FT NON_TER 4
 SQ SEQUENCE 4 AA; 420 MW; 6DD3DD6F00000000 CRC64;
 Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 Db 2 A 2
 RESULT 7
 EOSI_HUMAN STANDARD; PRT; 4 AA.
 ID EOSI_HUMAN
 AC P02731;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last annotation update)
 DE Eosinophilotoxic peptides.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=76078412; PubMed=1060093;
 RA Goetzl E.J., Austen K.F.;
 RT "Purification and synthesis of eosinophilotoxic tetrapeptides of
 RT human lung tissue: Identification as eosinophil chemotactic factor of
 RT anaphylaxis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
 CC -1- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG
 CC (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS
 CC (ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING
 CC EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE
 CC OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.
 DR GO; GO:0030105; P:anaphylaxis; IDA.
 DR GO; GO:0006935; P:chemotaxis; IDA.
 FT VARIANT 1
 V -> A (IN OTHER PEPTIDE).
 FT /FTID=VAR 005201.
 SQ SEQUENCE 4 AA; 390 MW; 6B05B862A0000000 CRC64;
 Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 Db 3 S 3
 RESULT 6
 DCMS_PSECH STANDARD; PRT; 4 AA.
 ID DCMS_PSECH
 AC P19918;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO

DE dehydrogenase subunit S) (CO-DH S) (Fragment).
 GN CUTS.
 OS Pseudomonas carboxydohydrogena.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae.
 OX NCBI_TaxID=290;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=90055678; PubMed=2818128;
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
 RT "Homology and distribution of CO dehydrogenase structural genes in
 RT carboxydohydrogenic bacteria.";
 RL Arch. Microbiol. 152:335-341(1989).
 CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
 CC dioxide.
 CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
 CC acceptor.
 CC -1- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.
 CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
 CC SMALL.
 DR PIR; P10146; P10146.
 KW Oxidoreductase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S.
 FT NON_TER 4
 SQ SEQUENCE 4 AA; 420 MW; 6DD3DD6F00000000 CRC64;
 Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 Db 2 A 2

RESULT 7
 EOSI_HUMAN STANDARD; PRT; 4 AA.
 ID EOSI_HUMAN
 AC P02731;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUL-1986 (Rel. 01, Last annotation update)
 DE Eosinophilotoxic peptides.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=76078412; PubMed=1060093;
 RA Goetzl E.J., Austen K.F.;
 RT "Purification and synthesis of eosinophilotoxic tetrapeptides of
 RT human lung tissue: Identification as eosinophil chemotactic factor of
 RT anaphylaxis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
 CC -1- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG
 CC (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS
 CC (ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING
 CC EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE
 CC OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.
 DR GO; GO:0030105; P:anaphylaxis; IDA.
 DR GO; GO:0006935; P:chemotaxis; IDA.
 FT VARIANT 1
 V -> A (IN OTHER PEPTIDE).
 FT /FTID=VAR 005201.
 SQ SEQUENCE 4 AA; 390 MW; 6B05B862A0000000 CRC64;
 Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 X 1
 Db 3 S 3

RESULT 8

FAR3_HIRME STANDARD; PRT; 4 AA.
 AC P42562;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE FMRamide-like neuropeptide YNRF-amide.
 OS Hirudo medicinalis (Medicinal leech).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
 OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.
 OX NCBI_TaxID=6421;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92195954; PubMed=1686933;
 RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
 RT "Identification of Rfamide neuropeptides in the medicinal leech.";
 RL Peptides 12:897-908(1991).
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE) FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD RES 4 4 AMIDATION.
 SQ SEQUENCE 4 AA; 598 MW; 69D4073B30000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 Y 1

RESULT 9

FAR4_HIRME STANDARD; PRT; 4 AA.
 AC P42563;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE FMRamide-like neuropeptide YNRF-amide.
 OS Hirudo medicinalis (Medicinal leech).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
 OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.
 OX NCBI_TaxID=6421;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92195954; PubMed=1686933;
 RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
 RT "Identification of Rfamide neuropeptides in the medicinal leech.";
 RL Peptides 12:897-908(1991).
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE) FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD RES 4 4 AMIDATION.
 SQ SEQUENCE 4 AA; 616 MW; 69D4068B300000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 Y 1

RESULT 10

FFKA_ANTEL STANDARD; PRT; 4 AA.
 AC P58705;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Antho-KAamide.
 OS Anthopleura elegantissima (Sea anemone).
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
 OC Nynanthaeae; Actiniidae; Anthopleura.
 OX NCBI_TaxID=6110;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92028852; PubMed=1681803;
 RA Nothacker H.-P., Rinehart K.L. Jr., Grimmelikhuijzen C.J.P.;
 RT "Isolation of L-3-phenyllactyl-Phe-Lys-Ala-NH2 (Antho-KAamide), a novel neuropeptide from sea anemones.";
 RL Biochem. Biophys. Res. Commun. 179:1205-1211(1991).
 RN [2]
 RP FUNCTION.
 RX MEDLINE=93391436; PubMed=8397415;
 RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
 RT "The expansion behaviour of sea anemones may be coordinated by two inhibitory neuropeptides, Antho-KAamide and Antho-Riamide.";
 RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
 CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle groups. May be involved in the expansion phase of feeding behaviour in sea anemones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC PIR; JQ1273; JQ1273.
 KW Neuropeptide; Amidation.
 FT MOD RES 1 1 L-3-PHENYLLACTYL.
 FT MOD RES 4 4 AMIDATION.
 SQ SEQUENCE 4 AA; 512 MW; 6DD339C9A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 4 A 4

RESULT 11

FLRF_HIRME STANDARD; PRT; 4 AA.
 AC P42561;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE FLRFamide.
 OS Hirudo medicinalis (Medicinal leech), and
 OS Helisoma trivolvis (Snail).
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
 OC Arynchobdellida; Hirudiniiformes; Hirudinidae; Hirudo.
 OX NCBI_TaxID=6421, 27815;
 RN [1]
 RP SEQUENCE.
 RX SPECIES=H.medicalinalis;
 RX MEDLINE=92195954; PubMed=1686933;
 RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
 RT "Identification of Rfamide neuropeptides in the medicinal leech.";
 RL Peptides 12:897-908(1991).
 RN [2]
 RP SEQUENCE.
 RX SPECIES=H.trivolvis; TISSUE=Kidney;
 RX MEDLINE=94286417; PubMed=7912428;
 RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
 RT "FMRFamide-related peptides from the kidney of the snail, Helisoma trivolvis.";
 RL Peptides 15:31-36(1994).
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE) FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD RES 4 4 AMIDATION.
 SQ SEQUENCE 4 AA; 582 MW; 69D40729A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 0; Gaps 0;

QY 1 X 1
 DB 1 F 1

RESULT 12

FLRN ANTEL STANDARD; PRT; 4 AA.
 AC P58707;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Antho-RNamide.
 OS Anthopleura elegantissima (Sea anemone).
 OC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Actiniaria;
 OC Nynantheae; Actiniidae; Anthopleura.
 OX NCBI_TaxID=6110;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RX MEDLINE=90319122; PubMed=1973541;
 RA Grimmlikhuijzen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,
 RA Reinscheid R.K., Nothacker H.-P., Staley A.L.;
 RT "Isolation of L-3-phenylacetyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea
 RT anemone neuropeptide containing an unusual amino-terminal blocking
 RT group.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414 (1990).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Neuron-specific.
 CC -!- MASS SPECTROMETRY: MW=549.3; METHOD=FAB.
 DR PIR; A35779; A35779.
 KW Neuropeptide; Amidation.
 FT MOD RES 1 1 L-3-PHENYLACTYL.
 FT MOD RES 4 4 AMIDATION.
 SQ SEQUENCE 4 AA; 549 MW; 64540729A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 F 1

RESULT 13

FMRF MACNI STANDARD; PRT; 4 AA.
 AC P01162;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE FMRamide (Peak C) (Cardioexcitatory neuropeptide).
 OS Macrocallista nimbosa (Sun-ray clam),
 OS Nereis virens (Sandworm),
 OS Hirudo medicinalis (Medicinal leech), and
 OS Helisoma trivolvis (Snail).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
 OC Veneroidea; Veneridae; Macrocallista.
 OX NCBI_TaxID=6594, 6353, 6421, 27815,
 RN [1]
 RP SEQUENCE, AND SYNTHESIS.
 RC SPECIES=M.nimbosa; TISSUE=Cerebral pedal, and Visceral ganglion;
 RX MEDLINE=77215956; PubMed=877582;
 RA Price D.A., Greenberg M.J.;
 RT "Structure of a molluscan cardioexcitatory neuropeptide.";
 RL Science 197:670-671 (1977).
 RN [2]
 RP SEQUENCE, AND CHARACTERIZATION.

RC SPECIES=M.nimbosa; TISSUE=Ganglion;
 RX MEDLINE=78012038; PubMed=909875;
 RA Price D.A., Greenberg M.J.;
 RT "Purification and characterization of a cardioexcitatory neuropeptide
 RT from the central ganglia of a bivalve mollusc.";
 RL Prep. Biochem. 7:261-281 (1977).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=N.virens;
 RX MEDLINE=90259866; PubMed=2342992;
 RA Krajniak K.G., Price D.A.;
 RT "Authentic FMRamide is present in the polychaete Nereis virens.";
 RL Peptides 11:75-77 (1990).
 RN [4]
 RP SEQUENCE.
 RC SPECIES=H.medicinalis;
 RX MEDLINE=92195954; PubMed=1686933;
 RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
 RT "Identification of RFamide neuropeptides in the medicinal leech.";
 RL Peptides 12:897-908 (1991).
 RN [5]
 RP SEQUENCE.
 RC SPECIES=H.trivolvis; TISSUE=Kidney;
 RX MEDLINE=94286417; PubMed=7912428;
 RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
 RT "FMRamide-related peptides from the kidney of the snail, Helisoma
 RT trivolvis";
 RL Peptides 15:31-36 (1994).
 CC -!- FUNCTION: MYOACTIVE; CARDIOEXCITATORY SUBSTANCE. PHARMACOLOGICAL
 CC ACTIVITIES INCLUDE AUGMENTATION, INDUCTION, AND REGULARIZATION OF
 CC CARDIAC CONTRACTION.
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)
 CC FAMILY.
 DR PIR; A01426; ECNK.
 DR PIR; A60418; A60418.
 KW Neuropeptide; Amidation.
 FT MOD RES 4 4 AMIDATION.
 SQ SEQUENCE 4 AA; 600 MW; 69D40699A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 F 1

RESULT 14

FYRI ANTEL STANDARD; PRT; 4 AA.
 AC P58706;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Antho-RNamide I (Contains: Antho-RNamide II).
 OS Anthopleura elegantissima (Sea anemone).
 OC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Actiniaria;
 OC Nynantheae; Actiniidae; Anthopleura.
 OX NCBI_TaxID=6110;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92270459; PubMed=1821096;
 RA Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,
 RA Grimmlikhuijzen C.J.P.;
 RT "Isolation of two novel
 RT biologically active L-3-phenylacetyl-Tyr-Arg-Ile-NH2 and its
 RT des-phenylacetyl fragment Tyr-Arg-Ile-NH2.";
 RL Peptides 12:1165-1173 (1991).
 RN [2]
 RP FUNCTION.
 RX MEDLINE=93391436; PubMed=8397415;
 RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmlikhuijzen C.J.P.;

RT "The expansion behaviour of sea anemones may be coordinated by two
 RT inhibitory neuropeptides, Antho-Ramide and Antho-Ramide.";
 RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
 CC -I- FUNCTION: Inhibits spontaneous contractions in several muscle
 CC groups. May be involved in the expansion phase of feeding
 CC behaviour in sea anemones.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- TISSUE SPECIFICITY: Neuron-specific.
 KW Neuropeptide; Amidation.
 FT CHAIN 1 4 ANTHO-RIAMIDE I.
 FT CHAIN 2 4 ANTHO-RIAMIDE II.
 FT MOD RES 1 1 L-3-PHENYLLACTYL.
 FT MOD RES 4 4 AMIDATION.
 SQ SEQUENCE 4 AA; 598 MW; 60441B59A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 F 1

RESULT 15

OCPL OCTMI STANDARD; PRT; 4 AA.
 AC P58648;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cardioactive peptides Ocp-1/Ocp-2.
 OS Octopus minor (Octopus)
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
 OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.
 OX NCBI_TaxID=89766;
 RN [1]
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=20336815; PubMed=10876044;
 RA Iwakoshi E., Hisada M., Minakata H.;
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,
 RT Octopus minor.";
 RL Peptides 21:623-630(2000).
 CC -I- FUNCTION: Cardioactive; has both positive chronotropic and
 CC inotropic effects on the heart. Ocp-2 is a 1000 time less
 CC active than Ocp-1.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- PTM: Ocp-2 has L-Phe instead of D-Phe.
 CC -I- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.
 KW Hormone; D-amino acid.
 FT MOD RES 2 2 D-PHENYLALANINE.
 SQ SEQUENCE 4 AA; 394 MW; 6AA879C810000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 G 1

Search completed: December 8, 2003, 09:14:14
 Job time : 2.05051 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 1.65657 Seconds
(without alignments)
232.212 Million cell updates/sec

Title: US-09-498-556C-85

Perfect score: 4
Sequence: 1 XXXX 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	3	3 RHTDTC	thyroliberin - Bom
2	0	0.0	3	3 RHPCT	thyroliberin - pig
3	0	0.0	3	3 RSHST	thyroliberin - she
4	0	0.0	3	3 A92971	thyroliberin - eas
5	0	0.0	3	3 GKHU	growth-modulating
6	0	0.0	3	3 A60898	bursin - chicken
7	0	0.0	3	3 A23751	spinal cord peptid
8	0	0.0	3	3 B23751	spinal cord peptid
9	0	0.0	3	3 A33802	thyrotropin-releas
10	0	0.0	3	3 A22565	R-phycoerythrin al
11	0	0.0	3	3 PQ0010	angiotensin-conver
12	0	0.0	3	3 S13894	histidinol dehydro
13	0	0.0	3	3 A43391	TRH-like tripeptid
14	0	0.0	3	3 E37196	bradykinin-potenti
15	0	0.0	3	3 F37196	bradykinin-potenti
16	0	0.0	3	3 I50412	gene p20K protein
17	0	0.0	3	3 PT0636	T-cell receptor be
18	0	0.0	3	3 PT0578	T-cell receptor be
19	0	0.0	3	3 PT0571	T-cell receptor be
20	0	0.0	3	3 PT0622	T-cell receptor be
21	0	0.0	3	3 I78890	tyrosine protein k
22	0	0.0	3	3 S68328	blood cell protein
23	0	0.0	3	3 S13892	cytochrome-c oxida
24	0	0.0	4	1 ECXAA	antho-Ramide neur
25	0	0.0	4	2 S18401	thyroglobulin - do
26	0	0.0	4	2 A02147	phagocytosis-stimu
27	0	0.0	4	2 A32039	tyrosine-melanocyt
28	0	0.0	4	2 ECNK	cardioexcitatory n
29	0	0.0	4	2 PL0140	carbon-monoxide de

30 0 0.0 4 2 PL0146 carbon-monoxide de
31 0 0.0 4 2 A37832 phenol 2-monooxyge
32 0 0.0 4 2 A48360 gamma subunit of p
33 0 0.0 4 2 I40697 biotin A - Citroba
34 0 0.0 4 2 A61300 22K superhelical D
35 0 0.0 4 2 I57745 D-mannosate hydrol
36 0 0.0 4 2 A41890 protein D - Escher
37 0 0.0 4 2 S43014 hypothetical prote
38 0 0.0 4 2 D41654 hypothetical prote
39 0 0.0 4 2 B43848 cell surface adhes
40 0 0.0 4 2 I40505 hypothetical prote
41 0 0.0 4 2 I40870 phospholipase C (E
42 0 0.0 4 2 T46627 endoglycanase F -
43 0 0.0 4 2 S53508 starvation-induced
44 0 0.0 4 2 S53508 ribosomal protein
45 0 0.0 4 2 S17255

ALIGNMENTS

RESULT 1

RHTDTC

thyroliberin - Bombina orientalis

C:Species: Bombina orientalis

C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C:Accession: A90919; A01415

R;Yasuhara, T.; Nakajima, T.

Chem. Pharm. Bull. 23, 3301-3303, 1975

A:Title: Occurrence of Pyr-His-Pro-NH-2 in the frog skin.

A:Reference number: A90919; MUID:76138399; PMID:815011

A:Accession: A90919

A:Molecule type: protein

A:Residues: 1-3 <IAS>

C:Superfamily: thyroliberin precursor

C:Keywords: amidated carboxyl end; cutaneous gland; hormone; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 Q 1

RESULT 2

RHPGT

thyroliberin - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001

C:Accession: A01415

R;Nair, R.M.G.; Barrett, J.F.; Bowers, C.Y.; Schally, A.V.

Biochemistry 9, 1103-1106, 1970

A:Title: Structure of porcine thyrotropin releasing hormone.

A:Reference number: A90560; MUID:70136150; PMID:4984938

A:Accession: A01415

A:Molecule type: protein

A:Residues: 1-3 <NAI>

R;Boiler, J.; Enzmann, F.; Folkers, K.; Bowers, C.Y.; Schally, A.V.

Biochem. Biophys. Res. Commun. 37, 705-710, 1969

A:Title: The identity of chemical and hormonal properties of the thyrotropin releasing

A:Reference number: A90167; MUID:70039904; PMID:4982117

A:Contents: annotation

A:Note: biological activities and Rf values (in 17 chromatographic systems) of the synt

C:Superfamily: thyroliberin precursor

C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;

Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 Q 1

RESULT 3
RHSHY
thyroliberin - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A93750; A01415
R;Desiderio Jr., D.M.; Burgess, R.; Dunn, T.F.; Vale, W.; Guillemin, R.; Ward, D.N.
Org. Mass Spectrom. 5, 221-228, 1971
A;Title: The elucidation of the primary structure of the hypothalamic thyroïd stimulating hormone
A;Reference number: A93750
A;Accession: A93750
A;Molecule type: protein
A;Residues: 1-3 <DES>
R;Burgus, R.; Dunn, T.F.; Desiderio, D.; Ward, D.N.; Vale, W.; Guillemin, R.
Nature 226, 321-325, 1970
A;Title: Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor.
A;Reference number: A93161; MUID:70163386; PMID:4985794
A;Contents: annotation
A;Note: physicochemical characteristics and biological activities of the natural and synthetic thyroliberin precursor
C;Superfamily: thyroliberin precursor
C;Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 Q 1

RESULT 4
thyroliberin - eastern newt (tentative sequence)
C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A92971; A01415
R;Grimm-Jorgensen, Y.; McKelvy, J.F.
J. Neurochem. 23, 471-478, 1974
A;Title: Biosynthesis of thyrotropin releasing factor by newt (Triturus viridescens) brain
A;Reference number: A92971; MUID:75035605; PMID:4214528
A;Accession: A92971
A;Molecule type: protein
A;Residues: 1-3 <GRI>
A;Note: a peptide with the chromatographic and electrophoretic characteristics of thyroliberin, or glutamic acid
C;Superfamily: thyroliberin precursor
C;Keywords: amidated carboxyl end; cutaneous gland; hormone; hypothalamus; pyroglutamic acid
F;1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 Q 1

RESULT 5
GKHU
growth-modulating peptide - human

C;Species: Homo sapiens (man)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A01421
R;Schlesinger, D.H.; Pickart, L.; Thaler, M.M.
Experientia 33, 324-325, 1977
A;Title: Growth-modulating serum tripeptide is glycyl-histidyl-lysine.
A;Reference number: A01421; MUID:77162369; PMID:858356
A;Accession: A01421
A;Molecule type: protein
A;Residues: 1-3 <SCH>
A;Note: this serum tripeptide is found to stimulate growth of some cell types and to inhibit growth of others
C;Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 G 1

RESULT 6
bursin - chicken
C;Species: Gallus gallus (chicken)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A60898
R;Audhya, T.; Kroon, D.; Heavner, G.; Viamontes, G.; Goldstein, G.
Science 231, 997-999, 1986
A;Title: Tripeptide structure of bursin, a selective B-cell-differentiating hormone of the bursa of Fabricius
A;Reference number: A60898; MUID:86122916; PMID:3484938
A;Accession: A60898
A;Molecule type: protein
A;Residues: 1-3 <AUD>
C;Superfamily: unassigned animal peptides
C;Keywords: amidated carboxyl end; hormone
F;3/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 K 1

RESULT 7
A23751
spinal cord peptide SCP-4 - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A23751
R;Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.
Arch. Biochem. Biophys. 240, 178-183, 1985
A;Reference number: A23751; MUID:85250425; PMID:4015098
A;Accession: A23751
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-3 <HSI>
C;Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 Q 1

RESULT 8

B23751
 spinal cord peptide SCP-5 - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
 C:Accession: B23751
 R:Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.
 Arch. Biochem. Biophys. 240, 178-183, 1985
 A:Reference number: A23751; MUID:85250425; PMID:4015098
 A:Accession: B23751
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-3 <HSI>
 C:Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 M 1

RESULT 9
 A33802
 thyrotropin-releasing hormone-like peptide - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
 C:Accession: A33802
 R:Cockle, S.M.; Aitken, A.; Beg, F.; Smyth, D.G.
 J. Biol. Chem. 264, 7788-7791, 1989
 A:Title: A novel peptide, pyroglutamylglutamyloproline amide, in the rabbit prostate comp
 A:Reference number: A33802; MUID:89255196; PMID:2498305
 A:Accession: A33802
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-3 <COC>
 C:Superfamily: unassigned animal peptides
 C:Keywords: amidated carboxyl end; pyroglutamic acid
 F.1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F.3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 Q 1

RESULT 10
 A22565
 R-phycoerythrin alpha-1 chain - red alga (Gastroclonium coulteri) (fragment)
 C:Species: Gastroclonium coulteri
 C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
 C:Accession: A22565
 R:Klotz, A.V.; Glazer, A.N.
 J. Biol. Chem. 260, 4856-4863, 1985
 A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.
 A:Reference number: A22565; MUID:85182601; PMID:3886644
 A:Accession: A22565
 A:Molecule type: protein
 A:Residues: 1-3 <KLO>

Query Match 0.0%; Score 0; DB 3; Length 3;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 2 Y 2

RESULT 11
 PQ0010
 angiotensin-converting enzyme inhibitor (FLP-3) - common fig
 N:Alternate names: ficus latex peptide 3
 C:Species: Ficus carica (common fig)
 C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
 C:Accession: PQ0010
 R:Maruyama, S.; Miyoshi, S.; Tanaka, H.
 Agric. Biol. Chem. 53, 2763-2767, 1989
 A:Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.
 A:Reference number: PQ0008
 A:Accession: PQ0010
 A:Molecule type: protein
 A:Residues: 1-3 <MAR>
 A:Experimental source: latex
 C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 0.0%; Score 0; DB 3; Length 3;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 L 1

RESULT 12
 S13894
 histidinol dehydrogenase (EC 1.1.1.23) - wild cabbage (fragment)
 C:Species: Brassica oleracea (wild cabbage)
 C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
 C:Accession: S13894
 R:Nagai, A.; Scheidegger, A.
 Arch. Biochem. Biophys. 284, 127-132, 1991
 A:Title: Purification and characterization of histidinol dehydrogenase from cabbage.
 A:Reference number: S13894; MUID:91112783; PMID:1989490
 A:Accession: S13894
 A:Molecule type: protein
 A:Residues: 1-3 <NAG>
 A:Experimental source: var. capitata
 C:Keywords: dimer; NAD; oxidoreductase

Query Match 0.0%; Score 0; DB 3; Length 3;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 X 1
 DB 1 X 1

RESULT 13
 A43391
 TRH-like tripeptide - alfalfa
 C:Species: Medicago sativa (alfalfa)
 C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
 C:Accession: A43391
 R:Lackey, D.B.
 J. Biol. Chem. 267, 17508-17511, 1992
 A:Title: Isolation and structural determination of a novel TRH-like tripeptide, pyroglu-
 A:Reference number: A43391; MUID:92388092; PMID:1517203
 A:Accession: A43391
 A:Molecule type: protein
 A:Residues: 1-3 <LAC>
 C:Keywords: amidated carboxyl end; pyroglutamic acid
 F.1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F.3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
 Best Local Similarity 0.0%; Pred. No. 0;
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 Q 1

RESULT 14

E37196
bradykinin-potentiating peptide 5 - island jararaca
C;Species: Bothrops insularis (island jararaca)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 30-Jun-2001
C;Accession: E37196
J;Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.
J. Protein Chem. 9, 221-227, 1990
A;Title: Primary structure and biological activity of bradykinin potentiating peptides
A;Reference number: A37196; MUID:90351557; PMID:2386615
A;Accession: E37196
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-3 <CIN>
C;Keywords: pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 Q 1

RESULT 15

F37196
bradykinin-potentiating peptide 6 - island jararaca
C;Species: Bothrops insularis (island jararaca)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: F37196
J;Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.
J. Protein Chem. 9, 221-227, 1990
A;Title: Primary structure and biological activity of bradykinin potentiating peptides
A;Reference number: A37196; MUID:90351557; PMID:2386615
A;Accession: F37196
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-3 <CIN>
C;Keywords: pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 Q 1

Search completed: December 8, 2003, 09:19:33
Job time : 2.65657 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:31 ; Search time 1.85859 Seconds
(without alignments)
91.060 Million cell updates/sec

Title: US-09-498-556C-357
Perfect score: 11
Sequence: 1 LXR4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	3	1	US-08-470-837-7
2	9	81.8	3	4	US-08-868-452-7
3	9	81.8	4	1	US-08-288-405A-4
4	9	81.8	4	1	US-08-336-343A-36
5	9	81.8	4	3	US-09-142-078-19
6	9	81.8	4	3	US-09-357-141-19
7	9	81.8	4	4	US-09-533-889-19
8	9	81.8	4	4	US-09-142-080-19
9	9	81.8	4	4	US-09-486-283C-4
10	9	81.8	5	1	US-07-626-923A-12
11	9	81.8	5	1	US-08-357-264-5
12	9	81.8	5	1	US-08-456-840-38
13	9	81.8	5	1	US-08-672-514-5
14	9	81.8	5	1	US-07-789-184-173
15	9	81.8	5	1	US-08-406-192-12
16	9	81.8	5	1	US-08-460-343B-37
17	9	81.8	5	1	US-08-460-343B-38
18	9	81.8	5	1	US-08-475-263-173
19	9	81.8	5	1	US-08-398-028B-37
20	9	81.8	5	1	US-08-398-028B-38
21	9	81.8	5	1	US-08-266-407A-38
22	9	81.8	5	1	US-08-485-886-173
23	9	81.8	5	2	US-08-504-265B-37
24	9	81.8	5	2	US-08-504-265B-38
25	9	81.8	5	2	US-08-545-151-12
26	9	81.8	5	2	US-08-392-973A-7
27	9	81.8	5	2	US-08-392-973A-8

28	9	81.8	5	2	US-08-392-973A-9
29	9	81.8	5	2	US-08-477-362-173
30	9	81.8	5	2	US-08-477-134-173
31	9	81.8	5	2	US-08-892-544-38
32	9	81.8	5	3	US-08-473-489A-173
33	9	81.8	5	3	US-08-485-695-173
34	9	81.8	5	3	US-08-018-760-173
35	9	81.8	5	3	US-09-246-500B-17
36	9	81.8	5	4	US-09-367-791A-108
37	9	81.8	5	6	5169933-16
38	9	81.8	5	6	5464756-42
39	9	81.8	6	1	US-08-236-427-11
40	9	81.8	6	1	US-07-923-724-35
41	9	81.8	6	2	US-08-631-427A-3
42	9	81.8	6	2	US-08-806-203-2
43	9	81.8	6	2	US-08-806-203-3
44	9	81.8	6	2	US-08-806-203-4
45	9	81.8	6	2	US-08-806-203-5

ALIGNMENTS

RESULT 1
US-08-470-837-7
; Sequence 7, Application US/08470837
; Patent No. 5800811
; GENERAL INFORMATION:
; APPLICANT: Nimmi, Marcel E.
; APPLICANT: Hall, Frederick L.
; APPLICANT: Tuan, Tai-Lan
; APPLICANT: Wu, Lingtao
; APPLICANT: Cheung, David T.
; TITLE OF INVENTION: Transforming Growth Factor B Fusion
; TITLE OF INVENTION: and
; TITLE OF INVENTION: Their Use in Wound Healing
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 11150 Santa Monica Boulevard, Suite 400
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90025-3395
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,837
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharp, Janice A.
; REGISTRATION NUMBER: 34,051
; REFERENCE/DOCKET NUMBER: 30630-1US01
; TELEPHONE: 310-445-1140
; TELEFAX: 310-445-9031
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-470-837-7

Query Match 81.8%; Score 9; DB 1; Length 3;
Best Local Similarity 66.7%; Pred. NO. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 LXR 3
  | |
Db 1 LTR 3

RESULT 2
US-08-868-452-7
; Sequence 7, Application US/08868452C
; Patent No. 6352972
; GENERAL INFORMATION:
; APPLICANT: Marcel E. Nimni
; APPLICANT: Frederick L. Hall
; APPLICANT: Lingtao Wu
; APPLICANT: Bo Han Shors
; APPLICANT: Edwin Shors
; TITLE OF INVENTION: BONE MORPHOGENETIC PROTEINS AND THEIR
; FILE OF INVENTION: USE IN BONE GROWTH
; FILE REFERENCE: 17972-11
; CURRENT APPLICATION NUMBER: US/08/868,452C
; CURRENT FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Human
US-08-868-452-7

Query Match 81.8%; Score 9; DB 4; Length 3;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
  | |
Db 1 LTR 3

RESULT 3
US-08-288-405A-4
; Sequence 4, Application US/08288405A
; Patent No. 5559009
; GENERAL INFORMATION:
; APPLICANT: Chandy, Kaniathara G.
; APPLICANT: Kalman, Katalin
; APPLICANT: Chandy, Grischa
; APPLICANT: Gutman, George A.
; TITLE OF INVENTION: A No. 5559009el Voltage-Gated Potassium Channel
; TITLE OF INVENTION: Gene
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert,
; ADDRESSEE: Attn: Walter H. Dreger
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/288,405A
; FILING DATE: 10-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/207,431
; FILING DATE: 04-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-59844-1/WHD

Qy 1 LXR 3
  | |
Db 1 LTR 3

RESULT 4
US-08-336-343A-36
; Sequence 36, Application US/08336343A
; Patent No. 5677144
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Alves, Frauke
; TITLE OF INVENTION: CK-2, A No. 5677144el Receptor Tyrosine Kinase
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,343A
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-065
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-336-343A-36

Query Match 81.8%; Score 9; DB 1; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
  | |
Db 2 LAR 4

RESULT 5
US-09-142-078-19
```

Sequence 19, Application US/09142078
Patent No. 6172041
GENERAL INFORMATION:
APPLICANT: McCabe, R. Tyler
APPLICANT: Zhou, Li-Ming
APPLICANT: Layer, Richard T.
TITLE OF INVENTION: Use of Conantokins
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
STREET: 555 Thirteenth Street, N.W., Suite 701-E
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/142,078
FILING DATE: 10-FEB-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO US97/12652
FILING DATE: 21-JUL-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/762,377
FILING DATE: 06-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/684,750
FILING DATE: 22-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 2314-135.A
TELEPHONE: 202-783-6040
TELEFAX: 202-783-6031
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4
OTHER INFORMATION: /note= "Xaa is
OTHER INFORMATION: gamma-carboxyglutamic acid"
US-09-142-078-19

Query Match 81.8%; Score 9; DB 3; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 1 LAR 3

RESULT 6
US-09-357-141-19
Sequence 19, Application US/09357141
Patent No. 6277825
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: McIntosh, J. Michael
APPLICANT: McCabe, R. Tyler
APPLICANT: Layer, Richard T.
APPLICANT: Zhou, Li-Ming

TITLE OF INVENTION: Use of Conantokins for Treating Pain
FILE REFERENCE: 2314-171
CURRENT APPLICATION NUMBER: US/09/357,141
CURRENT FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: US 09/283,277
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: US 09/142,078
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: WO US97/12652
PRIOR FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: US 08/762,377
PRIOR FILING DATE: 1996-12-06
PRIOR APPLICATION NUMBER: US 08/684,750
PRIOR FILING DATE: 1996-07-22
NUMBER OF SEQ ID NOS: 71
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 4
TYPE: PRT
ORGANISM: Conus radiatus
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (4)
OTHER INFORMATION: Xaa is gamma-carboxyglutamic acid.
US-09-357-141-19

Query Match 81.8%; Score 9; DB 3; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 1 LAR 3

RESULT 7
US-09-533-889-19
Sequence 19, Application US/09533889
Patent No. 639574
GENERAL INFORMATION:
APPLICANT: McCabe, R. Tyler
APPLICANT: Zhou, Li-Ming
APPLICANT: Layer, Richard T.
APPLICANT: Olivera, Baldomero M.
APPLICANT: McIntosh, J. Michael
TITLE OF INVENTION: Use of Conantokins
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
STREET: 555 Thirteenth Street, N.W., Suite 701-E
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/533,889
FILING DATE: 22 MAR-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/142,078
FILING DATE: 10-FEB-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO US97/12652
FILING DATE: 21-JUL-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/762,377
FILING DATE: 06-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/684,750

; FILING DATE: 22-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-168.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /note= "Xaa is
; gamma-carboxyglutamic acid"
; US-09-533-889-19

Query Match 81.8%; Score 9; DB 4; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
| |
Db 1 LAR 3

RESULT 8
US-09-142-080-19
; Sequence 19, Application US/09142080
; Patent No. 6515103
; GENERAL INFORMATION:
; APPLICANT: Abogadie, Fe C.
; Cruz, Lourdes J.
; Olivera, Baldomero M.
; Walker, Craig
; Colledge, Clark
; Hillyard, David R.
; Jimenez, Elsie
; Laver, Richard T.
; Zhou, Li-Ming
; McCabe, R. Tyler
; TITLE OF INVENTION: Conantokins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figs, Earnst & Manbeck, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701-E
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142,080
; FILING DATE: 11-May-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US97/12618
; FILING DATE: 21-JUL-1997
; APPLICATION NUMBER: US 08/684,742
; FILING DATE: 22-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-134.A

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /note= "Xaa is
; gamma-carboxyglutamic acid"
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
; US-09-142-080-19

Query Match 81.8%; Score 9; DB 4; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
| |
Db 1 LAR 3

RESULT 9
US-09-486-283C-4
; Sequence 4, Application US/09486283C
; Patent No. 6573243
; GENERAL INFORMATION:
; APPLICANT: ITAGAKI, Yasuhiro
; APPLICANT: KONNO, Katsuhiko
; APPLICANT: KAWAI, No. 6573243ufumi
; APPLICANT: TAKAYAMA, Hiroaki
; TITLE OF INVENTION: POMPILID WASP-DERIVED NEUROPEPTIDES
; FILE REFERENCE: 1830/48652
; CURRENT APPLICATION NUMBER: US/09/486,283C
; CURRENT FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: JP 9/241699
; PRIOR FILING DATE: 1997-08-25
; PRIOR APPLICATION NUMBER: PCT/JP98/03730
; PRIOR FILING DATE: 1998-08-24
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Pompilid sp.
; US-09-486-283C-4

Query Match 81.8%; Score 9; DB 4; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
| |
Db 1 LSR 3

RESULT 10
US-07-626-923A-12
; Sequence 12, Application US/07626923A
; GENERAL INFORMATION:
; APPLICANT: Yoshimura, Akihiko
; APPLICANT: Longmore, Gregory D.
; APPLICANT: Lodish, Harvey
; TITLE OF INVENTION: MUTANT EPO RECEPTOR AND USES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:

ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/626,923A
FILING DATE: 13 December 1990
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH190-08
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-626-923A-12

Query Match 81.8%; Score 9; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 1 LAR 3

RESULT 11
US-08-357-264-5
Sequence 5, Application US/08357264
Patent No. 5541077
GENERAL INFORMATION:
APPLICANT: BURNIE M., James P.
APPLICANT: MATTHEWS M., Ruth C.
TITLE OF INVENTION: FUNGAL STRESS PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman, Darby & Cushman
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,264
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/152669
FILING DATE: 16-NOV-1993
APPLICATION NUMBER: US 663897
FILING DATE: 14-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: KOKULIS, Paul N.
REGISTRATION NUMBER: 16773
REFERENCE/DOCKET NUMBER: PNK/5544/202253/DJP

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-357-264-5

Query Match 81.8%; Score 9; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 1 LSR 3

RESULT 12
US-08-456-840-38
Sequence 38, Application US/08456840
Patent No. 5597908
GENERAL INFORMATION:
APPLICANT: Taddei-Peters, W. C.
APPLICANT: Butler, Sandra M.
TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo No. 5597908el
STREET: 1330 Piccard Drive
CITY: Rockville
STATE: Maryland
COUNTRY: US
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,840
FILING DATE: 01-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/266,407
FILING DATE: 27-JUN-1994
APPLICATION NUMBER: US 08/172,461
FILING DATE: 21-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gormley, Mary E.
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-258-5200
TELEFAX: 301-977-0847
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-840-38

Query Match 81.8%; Score 9; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 1 LSR 3

RESULT 13
US-08-672-514-5
; Sequence 5, Application US/08672514
; Patent No. 5686248
; GENERAL INFORMATION:
; APPLICANT: BURNIE Mr., James P.
; APPLICANT: MATTHEWS Ms., Ruth C.
; TITLE OF INVENTION: FUNGAL STRESS PROTEINS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman, Darby & Cushman
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,514
; FILING DATE: 28-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/152,669
; FILING DATE: 16-NOV-1993
; APPLICATION NUMBER: US 08/152669
; FILING DATE: 16-NOV-1993
; APPLICATION NUMBER: US 663897
; FILING DATE: 14-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, Paul N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: PNK/5544/202253/DJP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-672-514-5

Query Match 81.8%; Score 9; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 1 LSR 3

RESULT 14
US-07-789-184-173
; Sequence 173, Application US/07789184
; Patent No. 5688768
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California

COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/789,184
FILING DATE: 19911107
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 22000-20502.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 34-0154

INFORMATION FOR SEQ ID NO: 173:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 5
OTHER INFORMATION: /note= "This position is R-NH2."
US-07-789-184-173

Query Match 81.8%; Score 9; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 3 LAR 5

RESULT 15
US-08-406-192-12
; Sequence 12, Application US/08406192
; Patent No. 5739287
; GENERAL INFORMATION:
; APPLICANT: Wilbur, D. Scott
; APPLICANT: Prathare, Pradip M
; TITLE OF INVENTION: Biotinylated Cobalamins
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; STREET: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: WA 98101-2333
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/406,192
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/224,831
; FILING DATE: 08-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Renzoni, George E
; REGISTRATION NUMBER: 37,919
; REFERENCE/DOCKET NUMBER: RECL18947
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 682 8100
; TELEFAX: (206) 224 0779
; TELEX: 4938023
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-406-192-12

Query Match 81.8%; Score 9; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.5e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 3 LAR 5

Search completed: December 8, 2003, 09:20:30
Job time : 2.85859 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:18:46 ; Search time 3.55556 Seconds
(without alignments)
209.232 Million cell updates/sec

Title: US-09-498-556C-357

Perfect score: 11
Sequence: 1 LXR 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	4	9 US-09-248-158-2	Sequence 2, Appli
2	9	81.8	4	10 US-09-818-656A-8	Sequence 8, Appli
3	9	81.8	4	11 US-09-852-910-5	Sequence 5, Appli
4	9	81.8	4	11 US-09-852-910-155	Sequence 155, App
5	9	81.8	4	12 US-10-357-467-19	Sequence 19, Appl
6	9	81.8	4	12 US-10-277-292-674	Sequence 674, App
7	9	81.8	4	12 US-10-280-340-674	Sequence 674, App
8	9	81.8	4	12 US-10-391-399-115	Sequence 115, App
9	9	81.8	4	12 US-10-391-399-116	Sequence 116, App
10	9	81.8	4	15 US-10-206-699-295	Sequence 295, App
11	9	81.8	4	15 US-10-028-075B-6	Sequence 6, Appli
12	9	81.8	4	15 US-10-029-206A-6	Sequence 6, Appli
13	9	81.8	4	16 US-10-244-709-6	Sequence 6, Appli
14	9	81.8	5	12 US-10-407-123-108	Sequence 108, App
15	9	81.8	5	12 US-10-302-817A-36	Sequence 36, Appl

16	9	81.8	5	14 US-10-006-557-7	Sequence 7, Appli
17	9	81.8	5	15 US-10-206-699-32	Sequence 32, Appl
18	9	81.8	6	10 US-09-486-734A-17	Sequence 17, Appl
19	9	81.8	6	10 US-09-486-734A-27	Sequence 27, Appl
20	9	81.8	6	10 US-09-770-102A-17	Sequence 17, Appl
21	9	81.8	6	11 US-09-866-512A-9	Sequence 9, Appli
22	9	81.8	6	11 US-09-500-700-73	Sequence 73, Appl
23	9	81.8	6	12 US-10-308-967-17	Sequence 17, Appl
24	9	81.8	6	12 US-09-373-182B-1	Sequence 1, Appli
25	9	81.8	6	12 US-10-131-616B-6	Sequence 6, Appli
26	9	81.8	6	12 US-10-277-292-693	Sequence 693, App
27	9	81.8	6	12 US-10-280-340-693	Sequence 693, App
28	9	81.8	6	15 US-10-163-198-80	Sequence 80, Appl
29	9	81.8	6	16 US-10-244-709-4	Sequence 4, Appli
30	9	81.8	7	9 US-09-248-158-1	Sequence 1, Appli
31	9	81.8	7	9 US-09-779-233-37	Sequence 37, Appl
32	9	81.8	7	9 US-09-779-233-38	Sequence 38, Appl
33	9	81.8	7	9 US-09-779-233-42	Sequence 42, Appl
34	9	81.8	7	9 US-09-779-233-43	Sequence 43, Appl
35	9	81.8	7	9 US-09-779-233-44	Sequence 44, Appl
36	9	81.8	7	9 US-09-796-264-26	Sequence 26, Appl
37	9	81.8	7	9 US-09-731-242A-42	Sequence 42, Appl
38	9	81.8	7	9 US-09-731-242A-43	Sequence 43, Appl
39	9	81.8	7	9 US-09-731-242A-45	Sequence 45, Appl
40	9	81.8	7	9 US-09-989-789-123	Sequence 123, App
41	9	81.8	7	9 US-09-989-789-124	Sequence 124, App
42	9	81.8	7	9 US-09-989-789-125	Sequence 125, App
43	9	81.8	7	9 US-09-989-789-126	Sequence 126, App
44	9	81.8	7	9 US-09-989-789-127	Sequence 127, App
45	9	81.8	7	9 US-09-989-789-129	Sequence 129, App

ALIGNMENTS

RESULT 1

US-09-248-158-2
; Sequence 2, Application US/09248158
; Patent No. US20020015678A1
; GENERAL INFORMATION:
; APPLICANT: Yuan, Zhengyu
; TITLE OF INVENTION: Direct Adsorption Scintillation Assay
; TITLE OF INVENTION: for Measuring Enzyme Activity and Assaying Biochemical
; TITLE OF INVENTION: Processes
; FILE REFERENCE: 342312000600
; CURRENT APPLICATION NUMBER: US/09/248,158
; PRIOR FILING DATE: 1999-02-09
; PRIOR FILING DATE: 60/074,272
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; US-09-248-158-2

Query Match 81.8%; Score 9; DB 9; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 2 LAR 4

RESULT 2

US-09-818-656A-8
; Sequence 8, Application US/09818656A
; Patent No. US20020142381A1

```
; GENERAL INFORMATION:
; APPLICANT: GONG, Fangcheng et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001191
; CURRENT APPLICATION NUMBER: US/09/818,656A
; CURRENT FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-818-656A-8

Query Match      81.8%; Score 9; DB 10; Length 4;
Best Local Similarity 66.7%; Pred. NO. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
      | |
Db      2 LTR 4

RESULT 3
US-09-852-910-5
; Sequence 5, Application US/09852910
; Publication No. US20030096297A1
; GENERAL INFORMATION:
; APPLICANT: Hamm, Heidi
; APPLICANT: Gilchrist, Annette
; TITLE OF INVENTION: Method For Identifying Inhibitors of G Protein Coupled Receptor S
; FILE REFERENCE: 2661-101
; CURRENT APPLICATION NUMBER: US/09/852,910
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/275,472
; PRIOR FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 271
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 4
; TYPE: PRT
; ORGANISM: mammal
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(4)
; OTHER INFORMATION: PAR-23
US-09-852-910-5

Query Match      81.8%; Score 9; DB 11; Length 4;
Best Local Similarity 66.7%; Pred. NO. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
      | |
Db      1 LSR 3

RESULT 4
US-09-852-910-155
; Sequence 155, Application US/09852910
; Publication No. US20030096297A1
; GENERAL INFORMATION:
; APPLICANT: Hamm, Heidi
; APPLICANT: Gilchrist, Annette
; TITLE OF INVENTION: Method For Identifying Inhibitors of G Protein Coupled Receptor S
; FILE REFERENCE: 2661-101
; CURRENT APPLICATION NUMBER: US/09/852,910
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US 60/275,472
; PRIOR FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 271
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(4)
; OTHER INFORMATION: G alpha t library linker sequence
US-09-852-910-155

Query Match      81.8%; Score 9; DB 11; Length 4;
Best Local Similarity 66.7%; Pred. NO. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 LXR 3
      | |
Db      2 LTR 4

RESULT 5
US-10-357-467-19
; Sequence 19, Application US/10357467
; Publication No. US20030194729A1
; GENERAL INFORMATION:
; APPLICANT: Abogadie, Fe C.
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Walker, Craig
; APPLICANT: Colledge, Clark
; APPLICANT: Hillyard, David R.
; APPLICANT: Jimenez, Elsie
; TITLE OF INVENTION: Conantokins
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESS: Rothwell, Figg, Ernst & Manbeck, p.c.
; STREET: 1425 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/357,467
; FILING DATE: 04-Feb-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/142,080
; FILING DATE: 15-MAY-2000
; APPLICATION NUMBER: WO US97/12618
; FILING DATE: 21-JUL-1997
; APPLICATION NUMBER: US 08/684,742
; FILING DATE: 22-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-256.A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Modified-site
```

```
;
; LOCATION: 4
; OTHER INFORMATION: /note="Xaa is
; gamma-carboxylglutamic acid"
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-357-467-19

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
   |
Db 1 LAR 3

RESULT 6
US-10-277-292-674
; Sequence 674, Application US/10277292
; Publication No. US20030199470A1
; GENERAL INFORMATION:
; APPLICANT: FARIS, MARY
; APPLICANT: HUBERT, RENE
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: APAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: CHALLITA-EID, PIA
; APPLICANT: JAKOBOVITZ, AVA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158PID7
; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND
; TITLE OF INVENTION: OTHER CANCERS
; FILE REFERENCE: 51158-20050.00
; CURRENT APPLICATION NUMBER: US/10/277,292
; CURRENT FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US/09/935,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,098
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 700
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 674
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-10-277-292-674

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
   |
Db 2 LTR 4

RESULT 7
US-10-280-340-674
; Sequence 674, Application US/10280340
; Publication No. US20030207835A1
; GENERAL INFORMATION:
; APPLICANT: FARIS, MARY
; APPLICANT: HUBERT, RENE
; APPLICANT: RAITANO, ARTHUR
; APPLICANT: APAR, DANIEL
; APPLICANT: LEVIN, ELANA
; APPLICANT: CHALLITA-EID, PIA
; APPLICANT: JAKOBOVITZ, AVA
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158PID7
; TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF BLADDER AND
; TITLE OF INVENTION: OTHER CANCERS
; FILE REFERENCE: 51158-20050.00
; CURRENT APPLICATION NUMBER: US/10/280,340
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US/09/935,430
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227,098
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 700
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 674
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide motif
US-10-277-292-674

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
   |
Db 2 LTR 4

RESULT 8
US-10-391-399-115
; Sequence 115, Application US/10391399
; Publication No. US20030219806A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Lora, Jose M.
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: NOVEL 18607, 15603, 69318, 12303, 48000,
; TITLE OF INVENTION: 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921
; TITLE OF INVENTION: AND 33751 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MPI03-0200NMIM
; CURRENT APPLICATION NUMBER: US/10/391,399
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/789,481
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: US 09/634,669
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: US 09/583,373
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: US 09/510,706
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 10/309,804
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: US 60/336,936
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 10/094,214
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US60/275,078
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/828,035
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
; Remaining Prior Application data removed -.See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

```
; OTHER INFORMATION: N-glycosylation site
US-10-391-399-115

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LSR 4

RESULT 9
US-10-391-399-116
; Sequence 116, Application US/10391399
; Publication No. US20030219806A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Gluckemann, Maria Alexandra
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Lora, Jose M.
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: NOVEL 18607, 15603, 69318, 12303, 48000,
; TITLE OF INVENTION: 52920, 5433, 38554, 57301, 58324, 55063, 52991, 59914, 59921
; TITLE OF INVENTION: AND 33751 MOLECULES AND USES THEREFOR
; FILE REFERENCE: MPI03-0200NMIM
; CURRENT APPLICATION NUMBER: US/10/391,399
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 09/789,481
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: US 09/634,669
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: US 09/583,373
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: US 09/510,706
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 10/309,804
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: US 60/336,936
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 10/094,214
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US60/275,078
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 09/828,035
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: N-glycosylation site
US-10-391-399-116

Query Match      81.8%; Score 9; DB 12; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LSR 4

RESULT 10
US-10-206-699-295
; Sequence 295, Application US/10206699
; Publication No. US20030100510A1

; GENERAL INFORMATION:
; APPLICANT: Sundaramoorthy, M.
; APPLICANT: Hudson, B.
; TITLE OF INVENTION: Crystallized structure of Type IV Collagen NCI Domain Hexamer
; FILE REFERENCE: MBHB 01-1017
; CURRENT APPLICATION NUMBER: US/10/206,699
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 60/308,523
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/351,289
; PRIOR FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/366,854
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/385,362
; PRIOR FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 295
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-206-699-295

Query Match      81.8%; Score 9; DB 15; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 1 LAR 3

RESULT 11
US-10-028-075B-6
; Sequence 6, Application US/10028075B
; Publication No. US20030113733A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Nisar A.
; APPLICANT: Benner, Robert
; TITLE OF INVENTION: Gene regulator
; FILE REFERENCE: 2183-5223US
; CURRENT APPLICATION NUMBER: US/10/028,075B
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EP 01203748.7
; PRIOR FILING DATE: 2001-10-04
; NUMBER OF SEQ ID NOS: 175
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: swiss/p36507/MPK2 Human
US-10-028-075B-6

Query Match      81.8%; Score 9; DB 15; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LAR 4

RESULT 12
US-10-029-206A-6
; Sequence 6, Application US/10029206A
; Publication No. US20030119720A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Nisar A.
; APPLICANT: Benner, Robert
; TITLE OF INVENTION: Oligopeptide treatment of anthrax
; FILE REFERENCE: 2183-5222US
```


;; CURRENT APPLICATION NUMBER: US/10/029,206A
;; CURRENT FILING DATE: 2001-12-21
;; PRIOR APPLICATION NUMBER: 09/821,380
;; PRIOR FILING DATE: 2001-03-29
;; NUMBER OF SEQ ID NOS: 175
;; SOFTWARE: Patentin ver. 2.1
;; SEQ ID NO 6
;; LENGTH: 4
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: peptide
US-10-029-206A-6

Query Match 81.8%; Score 9; DB 15; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
| |
Db 2 LAR 4

RESULT 13
US-10-244-709-6
;; Sequence 6, Application US/10244709
;; Publication No. US20030129769A1
;; GENERAL INFORMATION:
;; APPLICANT: FRED E. REGNIER
;; TITLE OF INVENTION: AFFINITY SELECTED SIGNATURE PEPTIDES FOR PROTEIN IDENTIFICATION
;; FILE REFERENCE: 290.0001 0103
;; CURRENT APPLICATION NUMBER: US/10/244,709
;; CURRENT FILING DATE: 2002-09-16
;; PRIOR APPLICATION NUMBER: 60/203,227
;; PRIOR FILING DATE: 2000-05-05
;; PRIOR APPLICATION NUMBER: 60/208,184
;; PRIOR FILING DATE: 2000-05-31
;; PRIOR APPLICATION NUMBER: 60/208,372
;; PRIOR FILING DATE: 2000-05-31
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: Patentin version 3.0
;; SEQ ID NO 6
;; LENGTH: 4
;; TYPE: PRT
;; ORGANISM: Artificial
;; FEATURE:
;; OTHER INFORMATION: C-terminal sequence
US-10-244-709-6

Query Match 81.8%; Score 9; DB 16; Length 4;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
| |
Db 2 LAR 4

RESULT 14
US-10-407-123-108
;; Sequence 108, Application US/10407123
;; Publication No. US20030181381A1
;; GENERAL INFORMATION:
;; APPLICANT: Himmelspach, Michele
;; Schlokot, Uwe
;; Dörner, Friedrich
;; Fisch, Andreas
;; Eibl, Johann
;; TITLE OF INVENTION: Factor X Analogues With
;; a Modified Protease Cleavage Site
;; NUMBER OF SEQUENCES: 122
;; CORRESPONDENCE ADDRESS:

;; ADDRESSER: Townsend and Townsend and Crew LLP
;; STREET: Two Embarcadero Center, Eighth Floor
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94111-3834
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/407,123
;; FILING DATE: 04-Apr-2003
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/09/367,791A
;; FILING DATE: 12-NO. US20030181381A1-1999
;; APPLICATION NUMBER: AT A 335/97
;; FILING DATE: 27-FEB-1997
;; APPLICATION NUMBER: WO PCT/AT98/00045
;; FILING DATE: 27-FEB-1998
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Aussenhus, Scott L.
;; REGISTRATION NUMBER: 42,471
;; REFERENCE/DOCKET NUMBER: 20695D-000700US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; TELEX: <Unknown>
;; INFORMATION FOR SEQ ID NO: 108:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 108:
US-10-407-123-108

Query Match 81.8%; Score 9; DB 12; Length 5;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
| |
Db 2 LTR 4

RESULT 15
US-10-302-817A-36
;; Sequence 36, Application US/10302817A
;; Publication No. US20030198978A1
;; GENERAL INFORMATION:
;; APPLICANT: ROZZELLE, James
;; APPLICANT: BOLCHAKOVA, Elena
;; TITLE OF INVENTION: THERMUS BROCKIANUS NUCLEIC ACID POLYMERASES
;; FILE REFERENCE: 4768US
;; CURRENT APPLICATION NUMBER: US/10/302,817A
;; CURRENT FILING DATE: 2002-11-22
;; PRIOR APPLICATION NUMBER: 60/334,434
;; PRIOR FILING DATE: 2001-11-30
;; NUMBER OF SEQ ID NOS: 58
;; SOFTWARE: Patentin version 3.2
;; SEQ ID NO 36
;; LENGTH: 5
;; TYPE: PRT
;; ORGANISM: Thermus brockianus
US-10-302-817A-36

Query Match 81.8%; Score 9; DB 12; Length 5;
Best Local Similarity 66.7%; Pred. No. 6.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
|
Db 1 LSR 3

Search completed: December 8, 2003, 09:34:04
Job time : 4.55556 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 5.45455 Seconds
(without alignments)
116.399 Million cell updates/sec

Title: US-09-498-556C-357
Perfect score: 11
Sequence: 1 LXR 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1980.DAT.*
2:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1981.DAT.*
3:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1982.DAT.*
4:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1983.DAT.*
5:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1984.DAT.*
6:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1985.DAT.*
7:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1986.DAT.*
8:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1987.DAT.*
9:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1988.DAT.*
10:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1989.DAT.*
11:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1990.DAT.*
12:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1991.DAT.*
13:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1992.DAT.*
14:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1993.DAT.*
15:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1994.DAT.*
16:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1995.DAT.*
17:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1996.DAT.*
18:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1997.DAT.*
19:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1998.DAT.*
20:	SIDS1/gcgdata/genseq/genseqp-emb1/AA1999.DAT.*
21:	SIDS1/gcgdata/genseq/genseqp-emb1/AA2000.DAT.*
22:	SIDS1/gcgdata/genseq/genseqp-emb1/AA2001.DAT.*
23:	SIDS1/gcgdata/genseq/genseqp-emb1/AA2002.DAT.*
24:	SIDS1/gcgdata/genseq/genseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Query	Score	Match	Length	DB	ID	Description
1		9	81.8	3	20	AAW84193	Peptide comprising
2		9	81.8	4	19	AAW48182	Conantokin peptide
3		9	81.8	4	19	AAW49964	Conantokin peptide
4		9	81.8	4	22	AAW79019	Amino acid sequenc
5		9	81.8	4	23	AAE24457	BoNT/A N-glycosyla
6		9	81.8	4	23	AAE20561	Soybean diverged d
7		9	81.8	4	23	AAE14417	C-terminal sequenc
8		9	81.8	4	23	AAW51953	PSA antibody prepa
9		9	81.8	4	23	AAW51957	PSA antibody prepa

10	9	81.8	4	24	ABJ36682	G protein coupled
11	9	81.8	4	24	ABJ36806	G protein coupled
12	9	81.8	4	24	ABU13761	Novel human transp
13	9	81.8	5	11	AAR03446	Accessory moiety d
14	9	81.8	5	15	AAR48999	Sequence of portio
15	9	81.8	5	15	AAR50135	Ovine growth hormo
16	9	81.8	5	15	AAR50136	Ovine growth hormo
17	9	81.8	5	15	AAR50134	Ovine growth hormo
18	9	81.8	5	16	AAR77333	Human apolipoprote
19	9	81.8	5	17	AAW09695	Cyclic pentapeptid
20	9	81.8	5	17	AAW09696	Cyclic pentapeptid
21	9	81.8	5	17	AAW00251	Subtilisin N62D/G1
22	9	81.8	5	17	AAW00252	Subtilisin N62D/G1
23	9	81.8	5	18	AAW37317	Mammalian haemoglo
24	9	81.8	5	21	AAW51458	AAV VP3 derived pe
25	9	81.8	5	22	AAE11109	Tryptic peptide #1
26	9	81.8	5	22	AAE11144	Chymotryptic pepti
27	9	81.8	5	22	AAE05461	MASP substrate #3.
28	9	81.8	5	23	AAE26308	Human RBPI protein
29	9	81.8	5	23	ABG77841	Targeting peptide
30	9	81.8	5	23	AAW51982	PSA antibody prepa
31	9	81.8	5	23	AAW51956	PSA antibody prepa
32	9	81.8	5	24	ABP55341	Cucurbita sp. pept
33	9	81.8	5	24	ABP55342	Spinacia oleracea
34	9	81.8	5	24	ABP75051	Proteome analysis
35	9	81.8	6	15	AAR46808	Phytase derived pe
36	9	81.8	6	15	AAR50141	Ovine growth hormo
37	9	81.8	6	15	AAR50142	Ovine growth hormo
38	9	81.8	6	15	AAR50143	Ovine growth hormo
39	9	81.8	6	15	AAR55741	dsRNA-dependent ki
40	9	81.8	6	17	AAR90511	Hybridoma ATCC HB-
41	9	81.8	6	18	AAW74602	3-methylcholanthre
42	9	81.8	6	18	AAW33665	Complement antagon
43	9	81.8	6	18	AAW37334	Mammalian haemoglo
44	9	81.8	6	18	AAW26572	Peptide immunoreac
45	9	81.8	6	19	AAW76723	Pancratic polypep

ALIGNMENTS

RESULT 1

AAW84193
ID AAW84193 standard; peptide: 3 AA;

AAW84193:

AA
DT 25-MAR-1999 (first entry)

Peptide comprising a proteinase site.

Proteinase site; bone morphogenetic fusion protein; bone binding site;
KW
bone morphogenetic protein; transforming growth factor beta;
KW
active fragment; wound healing; bone growth.
KW

OS Unidentified.

PN WO9855137-A1.

PD 10-DEC-1998.

02-JUN-1998; 98WO-US11189.

PR 03-JUN-1997; 97US-0868452.

PA (HALL/) HALL F L.

PA (NIMN//) NIMNI M E

PA (SHOK//) SHOKS E C.
PA (WULL//) WU L.

PI Hall FL. Han B.

 \mathbb{Z}

DR WPI; 1999-059875/05.

XX New bone morphogenetic fusion proteins - comprising a purification

PT tag and a bone morphogenetic active fragment, used for enhancing

PT wound healing or bone growth

XX

PS Claim 8; Page 34; 64pp; English.

XX

CC The present peptide represents a proteinase site used in the creation

CC of the bone morphogenetic fusion proteins of the invention. The bone

CC morphogenetic fusion protein may contain some or all of the following

CC elements: a purification tag, a proteinase site, an ECM/bone binding

CC site, a second proteinase site, and a bone morphogenetic protein

CC active fragment. The fusion proteins of the invention also includes

CC proteins that have transforming growth factor beta active fragments

CC instead of bone morphogenetic protein active fragments. The bone

CC morphogenetic fusion proteins can be used for enhancing wound healing

CC or bone growth.

XX

XX Sequence 3 AA;

Query Match 81.8%; Score 9; DB 20; Length 3;

Best Local Similarity 66.7%; Pred. No. 9.2e+05;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

Db 1 LTR 3

RESULT 2

ID AAW48182 standard; peptide; 4 AA.

XX

AC AAW48182;

XX

DT 30-JUN-1998 (first entry)

XX

DE Conantokin peptide derivative.

XX

XX Conantokin; predatory cone snail; treatment; neurologic disorder;

KW psychiatric disorder; anticonvulsant; neuroprotective; analgesic.

KW HIV infection; ophthalmic indication; memory; learning defect;

KW cognitive defect.

XX

XX Synthetic.

OS

XX Key Location/Qualifiers

FT Modified-site 4

FT /note= "gamma-carboxyglutamic acid"

XX

FN WO9803541-A1.

XX

XX 29-JAN-1998.

XX

XX 21-JUL-1997; 97WO-US12618.

XX

XX 22-JUL-1996; 96US-0684742.

XX

XX (COGN-) COGNETIX INC.

PA (UTAH) UNIV UTAH RES FOUND.

XX

XX Abogadie FC, Colledge C, Cruz LJ, Hillyard DR, Jimenez E;

PI Layer RT, McCabe RT, Olivera BM, Rivier JE, Shen GS;

PI Walker C, Zhou L;

XX

DR WPI; 1998-120694/11.

XX

XX New conantokin peptide(s) - useful for e.g. treating neurologic or

PT psychiatric disorders, or the management of pain

XX

PS Claim 15; Page 98; 122pp; English.

XX

CC The present sequence is a conantokin peptide derivative, which can

CC be used to treat neurologic and psychiatric disorders, e.g. as an

CC anticonvulsant, neuroprotective or analgesic agent. Neurologic and

CC psychiatric disorders include epilepsy, convulsions, neurotoxic

CC injury (associated with conditions of hypoxia, anoxia or ischaemia,

CC which typically follow stroke, cerebrovascular accident, brain or

CC spinal cord trauma, myocardial infarct, physical trauma, drowning,

CC suffocation, perinatal asphyxia or hypoglycaemic events),

CC neurodegeneration (associated with Alzheimer's disease, senile

CC dementia, Amyotrophic Lateral Sclerosis, Multiple Sclerosis,

CC Parkinson's disease, Huntington's disease, Down's Syndrome,

CC Korsakoff's disease, schizophrenia, AIDS dementia, multi-infarct

CC dementia, Binswanger dementia and neuronal damage associated with

CC uncontrolled seizures), chemical toxicity (such as addiction, and

CC morphine, opiate, opioid and barbiturate tolerance), pain (acute,

CC chronic, migraine), anxiety, major depression, manic-depressive

CC illness, obsessive-compulsive disorder, schizophrenia and mood

CC disorders (such as bipolar disorder, unipolar depression, dysthymia

CC and seasonal affective disorder) and dystonia (movement disorder),

CC sleep disorder, muscle relaxation and urinary incontinence. The

CC peptide can also be used to treat HIV infection, ophthalmic

CC indication and memory, learning or cognitive defects.

XX

XX Sequence 4 AA;

Query Match 81.8%; Score 9; DB 19; Length 4;

Best Local Similarity 66.7%; Pred. No. 9.2e+05;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3

Db 1 LAR 3

RESULT 3

AAW49964

ID AAW49964 standard; peptide; 4 AA.

XX

AC AAW49964;

XX

DT 30-JUN-1998 (first entry)

XX

DE Conantokin peptide derivative.

XX

XX Conantokin; predatory cone snail; treatment; neurologic disorder;

KW psychiatric disorder; anticonvulsant; neuroprotective; analgesic.

KW HIV infection; ophthalmic indication; memory; learning defect;

KW cognitive defect.

XX

XX Synthetic.

OS

XX Key Location/Qualifiers

FT Modified-site 4

FT /note= "gamma-carboxyglutamic acid"

XX

FN WO9803189-A1.

XX

XX 29-JAN-1998.

XX

XX 21-JUL-1997; 97WO-US12652.

XX

XX 06-DEC-1996; 96US-0762377.

PR 22-JUL-1996; 96US-0684750.

XX

XX (COGN-) COGNETIX INC.

PA

XX Layer RT, McCabe RT, Zhou L;

PI

XX WPI; 1998-120469/11.

DR

XX

XX Use of conantokin peptide(s) - for treating disorders involving

PT excessive excitation of nerve cells by excitatory amino acids or

PT agonists of the N-methyl-D-aspartate receptor

Claim 27; Page 69; 122pp; English.

The present sequence is a conantokin peptide derivative, which can be used to treat neurologic and psychiatric disorders, e.g. as an anticonvulsant, neuroprotective or analgesic agent. Neurologic and psychiatric disorders include epilepsy, convulsions, neurotoxic injury (associated with conditions of hypoxia, anoxia or ischaemia, which typically follow stroke, cerebrovascular accident, brain or spinal cord trauma, myocardial infarct, physical trauma, drowning, suffocation, perinatal asphyxia or hypoglycaemic events), neurodegeneration (associated with Alzheimer's disease, senile dementia, Amyotrophic Lateral Sclerosis, Multiple Sclerosis, Parkinson's disease, Huntington's disease, Down's Syndrome, Korsakoff's disease, schizophrenia, AIDS dementia, multi-infarct dementia, Binswanger dementia and neuronal damage associated with uncontrolled seizures), chemical toxicity (such as addiction, and morphine, opiate, opioid and barbiturate tolerance), pain (acute, chronic, migraine), anxiety, major depression, manic-depressive illness, obsessive-compulsive disorder, schizophrenia and mood disorders (such as bipolar disorder, unipolar depression, dysthymia and seasonal affective disorder) and dystonia (movement disorder), sleep disorder, muscle relaxation and urinary incontinence. The peptide can also be used to treat HIV infection, ophthalmic indication and memory, learning or cognitive defects.

XX SQ Sequence 4 AA;

Query Match 81.8%; Score 9; DB 19; Length 4;
 Best Local Similarity 66.7%; Pred. No. 9.2e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
 | |
 Db 1 LAR 3

RESULT 4
 AAG79019
 ID AAG79019 standard; peptide; 4 AA.
 XX AC AAG79019;
 XX DT 10-DEC-2001 (first entry)
 XX DE Amino acid sequence of conantokin R domain III.
 XX KW Conantokin; cone snail; nerve cell excitation; NMDA receptor; epilepsy;
 KW N-methyl-D-aspartate receptor; pain; psychiatric disorder;
 KW neurotoxic injury; hypoxia; anoxia; ischemia; neurodegeneration;
 KW chemical toxicity; addiction; drug craving; psychiatric disorder;
 KW anxiety; depression; obsessive compulsive disorder; schizophrenia;
 KW mood disorder; ophthalmic disorder; neurological disorder; dystonia;
 KW sleep disorder; muscle relaxation; urinary incontinence;
 KW cognition enhancement; HIV infection.
 XX OS Conus radiatus.
 XX FH Key Location/Qualifiers
 FT Modified-site 4 /note= "gamma-carboxyglutamic acid"
 FT US6277825-B1.
 XX PN 21-AUG-2001.
 XX PD 20-JUL-1999; 99US-0357141.
 XX PF 22-JUL-1996; 96US-0684750.
 XX PR 06-DEC-1996; 96US-0762377.
 XX PR 21-JUL-1997; 97WO-US12652.
 XX PR 10-FEB-1999; 99US-0142076.
 XX PR 01-APR-1999; 99US-0283277.

(UTAH) UNIV UTAH RES FOUND.
 (COGN-) COGNETIX INC.

Olivera BW, McIntosh JM, McCabe RT, Layer RT, Zhou L;
 WPT; 2001-601377/68.

Use of conantokin peptide or its derivatives or a conantokin peptide chimera for treating disorders e.g. migraine -

Claim 9; Column 80; 60pp; English.

AAG79012-43 and AAG790054-56 represent domains of conantokin peptides. Conantokin differ from conotoxins, in that they contain gamma-carboxyglutamic acid. The conantokin are derived from the venom of cone snails. They are used for the treatment of disorders in which the pathophysiology involves excessive excitation of nerve cells by excitatory amino acids or agonist of N-methyl-D-aspartate (NMDA) receptor. The conantokin peptides are used for the treatment of disorders such as pain; neurologic or psychiatric disorders such as epilepsy; for reducing neurotoxic injury associated with conditions of hypoxia, anoxia or ischemia; for treating neurodegeneration; for treating chemical toxicity such as addiction, drug craving, alcohol abuse, morphine, opioid and barbiturate tolerance; for treating psychiatric disorders such as anxiety, major depression, manic-depression, illness, obsessive compulsive disorder, schizophrenia or mood disorder; for treating ophthalmic disorder; for treating additional neurological disorders e.g. dystonia, sleep disorder, muscle relaxation and urinary incontinence; for memory/cognition enhancement; for treating HIV infection.

XX SQ Sequence 4 AA;

Query Match 81.8%; Score 9; DB 22; Length 4;
 Best Local Similarity 66.7%; Pred. No. 9.2e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
 | |
 Db 1 LAR 3

RESULT 5
 AAE24457
 ID AAE24457 standard; peptide; 4 AA.
 XX AC AAE24457;
 XX DT 04-OCT-2002 (first entry)
 XX DE BONT/A N-glycosylation site #1.
 XX KW Neurotoxin; biological persistence; blepharospasm; pain; therapy;
 KW neuromuscular disorder; cervical dystonia; oromandibular dystonia;
 KW spasmodic dysphonia; laryngeal dystonia; muscular tension; asthma;
 KW neuralgia; autonomic nervous system disorder; sweating; salivation;
 KW headache; neuropathy; botulinum toxin serotype A; BONT/A.
 XX OS Clostridium botulinum.
 XX FT WO200240506-A2.
 XX PN 23-MAY-2002.
 XX PD 16-NOV-2001; 2001WO-US44030.
 XX PF 17-NOV-2000; 2000US-249540P.
 XX PR (ALLR) ALLERGAN SALES INC.
 XX PA Steward LE, Spanoyannis A, Lin W, Aoki KR;

DR WPI; 2002-479904/51.
 XX Modified neurotoxin especially Clostridial toxins, useful for treating
 PT neuromuscular and autonomic nervous system disorder and pain, comprises
 PT structural modification to alter biological persistence of neurotoxin
 PT -
 XX
 XX Disclosure; Page 19; 55pp; English.
 XX
 CC The invention related to modified neurotoxins especially
 CC Clostridial botulinum toxins with altered biological persistence.
 CC These toxins comprise a structural modification which is effective to
 CC alter the biological persistence. Modified neurotoxins of the invention
 CC are used for treating biological disorders which include neuromuscular
 CC disorder e.g. strabismus, blepharospasm, spasmodic torticollis (cervical
 CC dystonia), oromandibular dystonia and spasmodic dysphonia (laryngeal
 CC dystonia), autonomic nervous system disorders e.g. excessive salivation
 CC and sweating, asthma etc. and pain e.g. headache, muscular tension,
 CC neuralgia and neuropathy. The present sequence is botulinum toxin
 CC serotype A (BoNT/A) N-glycosylation site. This peptide is derived from
 CC C. botulinum.
 XX
 SQ Sequence 4 AA;
 Query Match 81.8%; Score 9; DB 23; Length 4;
 Best Local Similarity 66.7%; Pred. No. 9.2e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LXR 3
 DB 2 LTR 4
 DE
 DE Soybean diverged delta-9 fatty acid desaturase peptide #7.
 XX
 XX Soybean; diverged delta-9 fatty acid desaturase; enzyme; cellular lipid;
 KW oil; transgenic plant; gene mapping; immunisation.
 XX
 OS Glycine max.
 XX
 PN WO200216565-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 22-AUG-2001; 2001WO-US26246.
 XX
 XX 22-AUG-2000; 2000US-226996P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Booth JR, Cahoon RE, Hitz WD, Kinney AJ, Yadav NS;
 DR WPI; 2002-269353/31.
 XX
 XX New delta-9 fatty acid desaturase polypeptides and polynucleotides,
 PT useful in creating transgenic plants having altered levels of mono-,
 PT poly- and unsaturated fatty acids and in increasing the unsaturation
 PT levels in cellular lipids -
 XX
 PS Claim 12; Page 48; 77pp; English.
 XX
 CC The present invention relates to diverged delta-9 fatty acid desaturase
 CC proteins and polynucleotides encoding such proteins. The nucleic acid
 CC sequences may be used to increase the level of unsaturation in cellular
 CC lipids, including oil, in tissues when the enzyme is absent or rate-

CC limiting, to isolate cDNAs and genes encoding homologous proteins from
 CC the same or other plant species and to create transgenic plants in which
 CC the polypeptides are present at higher or lower levels than normal or in
 CC cell types or developmental stages in which they are not normally found,
 CC thus altering the level of mono-, poly- and unsaturated fatty acids in
 CC those cells. They are useful as probes for genetic and physical gene
 CC mapping and as markers, e.g. restriction fragment length polymorphism
 CC (RFLP) markers. The peptides can be used to immunise animals to produce
 CC antibodies specific for the peptides and proteins. The present sequence
 CC is soybean diverged delta-9 fatty acid desaturase peptide (residues
 CC 256-259).
 XX
 SQ Sequence 4 AA;
 Query Match 81.8%; Score 9; DB 23; Length 4;
 Best Local Similarity 66.7%; Pred. No. 9.2e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LXR 3
 DB 2 LAR 4
 DE
 DE 26-MAR-2002 (first entry)
 XX
 DE C-terminal sequence of an arginine-containing peptide.
 XX
 KW Protein quantification; protein identification; signature peptide;
 KW mass spectrometric analysis; proteolytic peptide; affinity ligand;
 KW arginine; isobaric peptide.
 XX
 OS Unidentified.
 XX
 PN WO200186306-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 04-MAY-2001; 2001WO-US14418.
 XX
 PR 05-MAY-2000; 2000US-203227P.
 PR 31-MAY-2000; 2000US-208184P.
 PR 31-MAY-2000; 2000US-208372P.
 XX
 PA (PURD) PURDUE RES FOUND.
 PA (REGN/) REGNIER F E.
 PA (CHAK/) CHAKRABORTY A B.
 PA (DORM/) DORMADY S J.
 PA (GENG/) G'ENG M.
 PA (JIJJ/) JI J.
 PA (RIGG/) RIGGS L D.
 PA (SIOM/) SIOMA C S.
 PA (WANG/) WANG S.
 PA (ZHAN/) ZHANG X.
 XX
 PI Regnier FE, Chakraborty AB, Dormady SJ, G'eng M, Ji J, Riggs LD;
 PI Sioma CS, Wang S, Zhang X;
 DR WPI; 2002-089810/12.
 XX
 XX Analyzing differences in protein content in protein samples, useful for
 PT identifying protein(s) in a sample or multiple proteins in a single
 PT complex mixture, by employing mass spectrometric analysis of signature
 PT peptides -
 XX
 PS Example 9; Page 73; 106pp; English.
 XX
 CC The invention relates to analysing differences in protein content among

CC plural protein samples comprising employing mass spectrometric analysis
CC of proteolytic peptide fragments. The method comprises
CC fragmenting at least a first protein sample and a second protein
CC sample to produce a first peptide pool and a second peptide pool;
CC isotopically labelling at least a portion of the peptides in at
CC least one of the pools so as to permit resolution of otherwise
CC identical peptides in the first and second peptide pools by mass
CC analysis; contacting peptides from at least a portion of both of the
CC peptide pools with a capture moiety to yield affinity-selected peptides
CC comprising an affinity ligand, where the capture moiety selects for the
CC affinity ligand; and analysing the affinity-selected peptides by mass
CC spectrometry to determine one or more differences between the first and
CC second samples. The method of the invention is useful for identifying one
CC or more proteins in a sample, and is typically useful for identifying
CC multiple proteins in a single complex mixture. The method is also useful
CC for quantifying proteins in a sample or several samples. The advantages
CC of the novel method are that it is easier to separate peptides than
CC proteins; native structure of the protein does not have to be maintained
CC during the analysis; structural variants do not interfere; and
CC putative proteins suggested from DNA database can be recognised by
CC using a signature peptide probe. The present method permits
CC identification of a protein in a complex sample without purifying the
CC protein or obtaining its composite peptide signature. The present
CC sequence is C-terminal fragment of a differentially acetylated
CC arginine-containing peptide which is used to generate fragment ions for
CC distinguishing isobaric peptides.

SQ Sequence 4 AA;

Query Match 81.8%; Score 9; DB 23; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.2e+05;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
|
Db 2 LAR 4

RESULT 8

AAM51953
ID AAM51953 standard; peptide; 4 AA.

AC AAM51953;

DT 01-FEB-2002 (first entry)

DE PSA antibody preparation immunogen peptide #5.

XX PSA; prostate-specific antigen; prostatic carcinoma; pro-kallikrein 2;
KW human; antibody; immunogen.

OS Homo sapiens.

XX DE10032040-A1.

XX 29-NOV-2001.

XX 05-JUL-2000; 2000DE-1032040.

XX 24-MAY-2000; 2000DE-1025387.

XX (HOFF) ROCHE DIAGNOSTICS GMBH.

XX WPI; 2002-042633/06.

XX Antibody comprising specificity for some truncated forms of
PT pro-prostate specific antigen, useful in immunoassays for differential
PT diagnosis of prostatic carcinoma -

PS Disclosure; Page 11; 16pp; German.

XX The present invention relates to an antibody specific for (-5,-6 and
CC -7)-proPSA (prostate-specific antigen), but with no reactivity with (-4)

CC or shorter forms of the protein. The antibody can be used to diagnose
CC prostatic carcinoma and to detect human pro-kallikrein 2. The present
CC sequence is a peptide which forms part of a peptide-containing immunogen
CC described in the exemplification of the invention.

SQ Sequence 4 AA;

Query Match 81.8%; Score 9; DB 23; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.2e+05;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
|
Db 2 LSR 4

RESULT 9

AAM51957
ID AAM51957 standard; peptide; 4 AA.

AC AAM51957;

DT 01-FEB-2002 (first entry)

XX PSA antibody preparation immunogen peptide #9.

XX PSA; prostate-specific antigen; prostatic carcinoma; pro-kallikrein 2;
KW human; antibody; immunogen.

OS Homo sapiens.

XX DE10032040-A1.

XX 29-NOV-2001.

XX 05-JUL-2000; 2000DE-1032040.

XX 24-MAY-2000; 2000DE-1025387.

XX (HOFF) ROCHE DIAGNOSTICS GMBH.

XX WPI; 2002-042633/06.

XX Antibody comprising specificity for some truncated forms of
PT pro-prostate specific antigen, useful in immunoassays for differential
PT diagnosis of prostatic carcinoma -

PS Disclosure; Page 11; 16pp; German.

XX The present invention relates to an antibody specific for (-5,-6 and
CC -7)-proPSA (prostate-specific antigen), but with no reactivity with (-4)
CC or shorter forms of the protein. The antibody can be used to diagnose
CC prostatic carcinoma and to detect human pro-kallikrein 2. The present
CC sequence is a peptide which forms part of a peptide-containing immunogen
CC described in the exemplification of the invention.

SQ Sequence 4 AA;

Query Match 81.8%; Score 9; DB 23; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.2e+05;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
|
Db 1 LSR 3

RESULT 10

ABJ36682
ID ABJ36682 standard; Peptide; 4 AA.

XX ABJ36682;

XX

```
DT 01-MAY-2003 (first entry)
DE G protein coupled receptor related peptide SEQ ID No 5.
XX
XX
XX Nootropic; cardiant; antiarteriosclerotic; hypotensive; cytostatic;
KW antibacterial; analgesic; antiasthmatic; antiinflammatory;
KW osteopathic; neuroprotective; anxiolytic; anorectic; lead compound;
KW G protein coupled receptor signaling inhibitor; GPCR; library;
KW high throughput screening assay; stroke; myocardial infarction;
KW restenosis; atherosclerosis; hypotension; cancer; infection; asthma;
KW septic shock; pain; allergic disorder; inflammatory bowel disease;
KW osteoporosis; obesity; psychotic; neurological disorder; anxiety;
KW schizophrenia; Alzheimer's disease.
XX Mammalia sp.
OS
XX WO200272778-A2.
XX
XX 19-SEP-2002.
XX
XX 14-MAR-2002; 2002WO-US07561.
XX
XX 14-MAR-2001; 2001US-275472P.
XX
XX 11-MAY-2001; 2001US-0852910.
XX
XX (CUEB-) CUE BIOTECH.
XX
XX Gilchrist A, Hamm HE;
XX
XX WPI; 2003-247841/24.
XX
XX Identifying G protein coupled receptor (GPCR) signaling inhibitors,
PT useful in screening drugs for treating stroke, cancers or pain, by
PT identifying compounds that block GPCR mediated signaling with high
PT affinity and specificity -
XX
XX Disclosure; Page 12; 94pp; English.
XX
XX The invention relates to a novel method for identifying a G protein
CC coupled receptor (GPCR) signaling inhibitor. The novel method comprises
CC selecting or identifying a member of a library of peptides and/or
CC candidate compounds, having binding to a GPCR of higher affinity than
CC that of the native peptide. The peptide library is based on a native GPCR
CC binding peptide. The method is useful for identifying inhibitors of a G
CC protein coupled receptor (GPCR) signaling. The method is particularly
CC useful for identifying drugs that antagonise the binding between a GPCR
CC and its extracellular ligand(s). The method is especially useful in
CC modern high throughput screening assays for identifying potent lead
CC compounds. The compounds, peptides or inhibitors identified by the method
CC are useful for preventing, ameliorating or treating diseases in which
CC GPCR signaling is a causative factor or in which a specific class of G
CC protein is relevant, e.g. stroke, myocardial infarction, restenosis,
CC atherosclerosis, hypotension, cancers, infections, septic shock, pain,
CC allergic disorders, asthma, inflammatory bowel disease, osteoporosis,
CC obesity, or psychotic and neurological disorders (e.g. anxiety,
CC schizophrenia or Alzheimer's disease). This sequence represents a peptide
CC relating to the G protein coupled receptors of the invention.
XX
XX Sequence 4 AA;
SQ
Query Match 81.8%; Score 9; DB 24; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.2e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LXR 3
| |
Db 1 LSR 3
| |
RESULT 11
ABJ36806
ID ABJ36806 standard; Peptide; 4 AA.
XX
```

```
AC ABJ36806;
XX
XX 01-MAY-2003 (first entry)
XX
XX G protein coupled receptor related peptide SEQ ID No 155.
DE
XX
XX Nootropic; cardiant; antiarteriosclerotic; hypotensive; cytostatic;
KW antibacterial; analgesic; antiasthmatic; antiinflammatory;
KW osteopathic; neuroprotective; anxiolytic; anorectic; lead compound;
KW G protein coupled receptor signaling inhibitor; GPCR; library;
KW high throughput screening assay; stroke; myocardial infarction;
KW restenosis; atherosclerosis; hypotension; cancer; infection; asthma;
KW septic shock; pain; allergic disorder; inflammatory bowel disease;
KW osteoporosis; obesity; psychotic; neurological disorder; anxiety;
KW schizophrenia; Alzheimer's disease.
XX Unidentified.
OS
XX WO200272778-A2.
XX
XX 19-SEP-2002.
XX
XX 14-MAR-2002; 2002WO-US07561.
XX
XX 14-MAR-2001; 2001US-275472P.
XX
XX 11-MAY-2001; 2001US-0852910.
XX
XX (CUEB-) CUE BIOTECH.
XX
XX Gilchrist A, Hamm HE;
XX
XX WPI; 2003-247841/24.
XX
XX Identifying G protein coupled receptor (GPCR) signaling inhibitors,
PT useful in screening drugs for treating stroke, cancers or pain, by
PT identifying compounds that block GPCR mediated signaling with high
PT affinity and specificity -
XX
XX Example 8; Page 178; 94pp; English.
XX
XX The invention relates to a novel method for identifying a G protein
CC coupled receptor (GPCR) signaling inhibitor. The novel method comprises
CC selecting or identifying a member of a library of peptides and/or
CC candidate compounds, having binding to a GPCR of higher affinity than
CC that of the native peptide. The peptide library is based on a native GPCR
CC binding peptide. The method is useful for identifying inhibitors of a G
CC protein coupled receptor (GPCR) signaling. The method is particularly
CC useful for identifying drugs that antagonise the binding between a GPCR
CC and its extracellular ligand(s). The method is especially useful in
CC modern high throughput screening assays for identifying potent lead
CC compounds. The compounds, peptides or inhibitors identified by the method
CC are useful for preventing, ameliorating or treating diseases in which
CC GPCR signaling is a causative factor or in which a specific class of G
CC protein is relevant, e.g. stroke, myocardial infarction, restenosis,
CC atherosclerosis, hypotension, cancers, infections, septic shock, pain,
CC allergic disorders, asthma, inflammatory bowel disease, osteoporosis,
CC obesity, or psychotic and neurological disorders (e.g. anxiety,
CC schizophrenia or Alzheimer's disease). This sequence represents a peptide
CC relating to the G protein coupled receptors of the invention.
XX
XX Sequence 4 AA;
SQ
Query Match 81.8%; Score 9; DB 24; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.2e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LXR 3
| |
Db 2 LTR 4
| |
RESULT 12
ABU13761
```


ID ABU13761 standard; Peptide; 4 AA.
XX AC ABU13761;
XX XX
DT 25-FEB-2003 (first entry)
XX DE
XX Novel human transporter protein related peptide #4.
XX Human; gamma-aminobutyric acid; GABA; neurotransmitter transporter;
KW transgenic animal; ribozyme design; drug screening; gene therapy.
XX OS
XX Homo sapiens.
XX XX
PN US2002142381-A1.
XX XX
PD 03-OCT-2002.
XX XX
PF 28-MAR-2001; 2001US-0818656.
XX XX
PR 28-MAR-2001; 2001US-0818656.
XX XX
PA (GONG/) GONG F.
PA (KETCHUM/) KETCHUM K A.
PA (DFRA/) DI FRANCESCO V.
PA (BEAS/) BEASLEY E M.
XX XX
PI Gong F, Ketchum KA, Di Francesco V, Beasley EM;
XX WPI; 2003-102517/09.
XX XX
PT Novel human transporter protein, related to gamma-aminobutyric acid
PT neurotransmitter transporter subfamily useful as model for developing
PT human therapeutic targets and serves as target for human therapeutics -
XX XX
PS Disclosure; Page 45; 114pp; English.
XX XX
CC The invention describes an isolated human transporter peptide (I) that
CC is related to the gamma-aminobutyric acid (GABA) neurotransmitter
CC transporter subfamily. (I) and the nucleic acid encoding it (II) can be
CC used as models for the development of human therapeutic targets, aid in
CC the identification of therapeutic proteins and serve as targets for the
CC development of human therapeutic agents that modulate transporter
CC activity. (I) is used to raise antibodies or to elicit another immune
CC response, as a reagent in assays designed to quantitatively determine
CC levels of the protein in biological fluids, and as markers for tissues
CC in which the corresponding protein is preferentially expressed. The
CC transporter proteins isolated from humans and their human/mammalian
CC orthologues serve as targets for identifying agents for use in mammalian
CC therapeutic applications, and biological assays related to transporter
CC proteins that are related to members of the GABA neurotransmitter
CC transporter subfamily. The proteins and peptides also provide a target
CC for diagnosing a disease or predisposition to disease mediated by the
CC peptide, and are useful for treating a disorder characterised by absence
CC of, inappropriate, unwanted or altered expression of the protein. The
CC antibodies are also useful for assessing normal and aberrant subcellular
CC localisation of cells in various tissues in an organism, in
CC pharmacogenomic analysis, for tissue typing and for inhibiting protein
CC function. (II) is useful for constructing recombinant vectors, host cells
CC and transgenic animals; for designing ribozymes; in drug screening;
CC in diagnostic assays for qualitative changes in gene expression;
CC particularly in qualitative changes that lead to pathology; in gene
CC therapy; and to detect mutations in genes encoding transporters. This
CC is the amino acid sequence of a polypeptide related to the novel human
CC aminobutyric acid (GABA) transporter related protein.
XX XX
SQ Sequence 4 AA;
Query Match 81.8%; Score 9; DB 24; Length 4;
Best Local Similarity 66.7%; Pred. No. 9.2e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LXR 3
| |

Db 2 LTR 4
RESULT 13
AAR03446
ID AAR03446 standard; protein; 5 AA.
XX AC AAR03446;
XX XX
DT 02-AUG-1990 (first entry)
XX XX
DE Accessory moiety derived from transferrin receptor.
XX KW Accessory moiety; transferrin receptor; tumour therapy.
XX OS Synthetic.
XX XX
PN EP359347-A.
XX XX
PD 21-MAR-1990.
XX XX
PF 14-AUG-1989; 89EP-0250014.
XX XX
PR 15-AUG-1988; 88US-0232337.
XX XX
PA (NEOR-) NEORX CORP.
XX XX
PI Anderson DC, Morgan AC, Abrams PG, Nichols EJ, Fritzberg AR;
XX WPI; 1990-085154/12.
XX XX
PT Covalently linked complex for tumour treatment - comprises
PT treating with protein, cytotoxic agent and enhancing moiety.
XX XX
PS Claim 13; Page 22; 23pp; English.
XX XX
CC The sequence is one of several possible accessory moieties which can be
CC attached to the C-terminus of an anchoring moiety which in turn is
CC linked to a targeting protein and a cytotoxic agent. When the anchoring
CC peptide intercalates into the target cell plasma membrane the accessory
CC moiety is translocated across the membrane and protrudes into the
CC cytoplasm. It can then be phosphorylated by cellular kinases at neutral
CC pH. This irreversibly anchors the complex to the membrane and may
CC enhance the translocation into the cytoplasm. The complex is useful for
CC treatment and diagnosis of tumours.
CC See also AAR03435-60.
XX XX
SQ Sequence 5 AA;
Query Match 81.8%; Score 9; DB 11; Length 5;
Best Local Similarity 66.7%; Pred. No. 9.2e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LXR 3
| |

Db 3 LAR 5
RESULT 14
AAR48999
ID AAR48999 standard; Protein; 5 AA.
XX AC AAR48999;
XX XX
DT 02-SEP-1994 (first entry)
XX XX
DE Sequence of portion of wild-type murine erythropoietin receptor
DE (nEPOR).
XX KW Erythropoietin receptor; EPO; ss.
XX XX
XX Mus musculus.

```

FH Key          Location/Qualifiers
FT Misc-difference 3
FT /note= "Cys in cEPOR"
XX
XX
XX
XX US5292654-A.
XX
XX 08-MAR-1994.
XX
XX 13-DEC-1990; 90US-0626923.
XX
XX 13-DEC-1990; 90US-0626923.
XX
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
XX Lodish H, Longmore GD, Yoshimura A;
XX
XX WPI; 1994-082328/10.
XX
XX N-PSDB; AAR48999.
XX
XX Mutant erythropoietin (EPO) receptor gene - hypersensitive to
XX EPO, useful in assay for identifying compounds, mimicking EPO
XX action
XX
XX Disclosure; Fig 1C; 13pp; English.
XX
XX nEPOR is wild-type EPOR. cEPOR is a constitutive EPOR. It contains
XX a point mutation (a transition from C to T at nucleotide 484) which
XX causes one substitution (Arg to Cys) in the exoplasmic domain of
XX EPOR at codon 129 of the predicted N-terminus. The Arg to Cys point
XX mutation in cEPOR is sufficient to induce factor-independent
XX growth.
XX
XX SQ Sequence 5 AA;
XX
XX Query Match 81.8%; Score 9; DB 15; Length 5;
XX Best Local Similarity 66.7%; Pred. No. 9.2e+05;
XX Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 LXR 3
XX |
XX 1 LAR 3
XX
XX RESULT 15
XX AAR50135
XX ID AAR50135 standard; peptide; 5 AA.
XX
XX AC AAR50135;
XX
XX DT 25-MAR-2003 (updated)
XX DT 23-SEP-1994 (first entry)
XX
XX DE Ovine growth hormone residues 92-96.
XX
XX KW Bovine; ovine; porcine; growth hormone; bGH; oGH; pGH; antigen;
XX growth hormone activity; potentiate; enhance; increase.
XX
XX OS Bos taurus.
XX OS Ovis aries.
XX OS Sus scrofa.
XX
XX PN W09405697-A1.
XX
XX PD 17-MAR-1994.
XX
XX PF 07-SEP-1993; 93WO-GB01887.
XX
XX PR 07-SEP-1992; 92GB-0018907.
XX PR 09-AUG-1993; 93GB-0016508.
XX
XX PA (AGRI-) AGRIC & FOOD RES COUNCIL.
XX PA (HANN-) HANNAH RES INST.
XX

```

```

PI Beattie J, Holder AT;
XX
XX WPI; 1994-101122/12.
XX
XX New antigenic molecules derived from growth hormone - and derived
XX antibodies, used to stimulate the activity of growth hormone
XX
XX Claim 4; Page 7; 64pp; English.
XX
XX This peptide corresponds to residues 92-96 of bovine, ovine and
XX porcine growth hormone. The peptide is rendered antigenic and is
XX then capable of stimulating the production of antibodies which,
XX when in the appropriate formulation, potentiate the effect of
XX growth hormone.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 5 AA;
XX
XX Query Match 81.8%; Score 9; DB 15; Length 5;
XX Best Local Similarity 66.7%; Pred. No. 9.2e+05;
XX Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 LXR 3
XX |
XX 2 LSR 4
XX
XX Search completed: December 8, 2003, 09:16:40
XX Job time : 6.45455 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 4.20202 Seconds
(without alignments)
245.646 Million cell updates/sec

Title: US-09-498-556C-357

Perfect score: 11
Sequence: 1 LXR 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	8	Q16468	Q16468 homo sapien
2	9	81.8	8	Q9T4Y2	Q9T4Y2 asterina pe
3	9	81.8	8	O89965	O89965 polyomaviru
4	9	81.8	9	Q9T688	Q9T688 gecko gecko
5	9	81.8	9	Q94VH4	Q94VH4 varanus gla
6	9	81.8	9	Q94VD8	Q94VD8 varanus nil
7	9	81.8	9	Q94V18	Q94V18 varanus ere
8	9	81.8	9	Q94VC6	Q94VC6 varanus pil
9	9	81.8	9	Q94VE1	Q94VE1 varanus mer
10	9	81.8	9	Q91BM8	Q91BM8 simian viru
11	9	81.8	12	Q9PYK1	Q9PYK1 simian viru
12	9	81.8	10	Q958K9	Q958K9 rana boylli
13	9	81.8	10	Q9TG86	Q9TG86 diploglossu
14	9	81.8	10	Q94V97	Q94V97 varanus spe
15	9	81.8	10	Q94VD5	Q94VD5 varanus oli
16	9	81.8	10	Q94VC9	Q94VC9 varanus pan

17	9	81.8	10	8	Q94VF0	Q94VF0 varanus kin
18	9	81.8	10	8	Q9TG47	Q9TG47 ophisaurus
19	9	81.8	10	8	P92771	P92771 xenosaurus
20	9	81.8	10	8	Q94V85	Q94V85 varanus var
21	9	81.8	10	8	Q9TG32	Q9TG32 ophisaurus
22	9	81.8	10	8	Q94PD8	Q94PD8 varanus sca
23	9	81.8	10	8	Q9TG38	Q9TG38 ophisaurus
24	9	81.8	10	8	P92576	P92576 bipes bipor
25	9	81.8	10	8	Q9TG35	Q9TG35 ophisaurus
26	9	81.8	10	8	Q94VD2	Q94VD2 varanus pan
27	9	81.8	10	8	Q9TFV5	Q9TFV5 eublepharus
28	9	81.8	10	8	Q8SHN1	Q8SHN1 bradypodion
29	9	81.8	11	3	Q9C1R7	Q9C1R7 saccharomyc
30	9	81.8	11	7	Q97900	Q97900 oreochromis
31	9	81.8	11	7	Q97917	Q97917 oreochromis
32	9	81.8	11	7	Q97902	Q97902 oreochromis
33	9	81.8	11	7	Q97921	Q97921 pseudotroph
34	9	81.8	11	7	Q97901	Q97901 oreochromis
35	9	81.8	11	7	Q97916	Q97916 oreochromis
36	9	81.8	11	7	Q97905	Q97905 oreochromis
37	9	81.8	11	7	Q97899	Q97899 oreochromis
38	9	81.8	11	7	Q978121	Q978121 oreochromis
39	9	81.8	11	7	Q97904	Q97904 oreochromis
40	9	81.8	11	7	Q97903	Q97903 oreochromis
41	9	81.8	11	8	Q94V94	Q94V94 varanus sto
42	9	81.8	11	8	Q9G631	Q9G631 calotes nig
43	9	81.8	11	8	Q9G622	Q9G622 salea horsf
44	9	81.8	11	8	Q94VG8	Q94VG8 varanus gou
45	9	81.8	11	8	Q9G365	Q9G365 calotes emm

ALIGNMENTS

RESULT 1

Q16468 PRELIMINARY; PRT; 8 AA.
AC Q16468;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DE 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DT DNA for cosmid CC13-1134 PCR primer 1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96435920; PubMed=8838806;
RA James L.A., Ogilvie D.J., Yamakawa K., Nakamura Y., Stirling C.J.,
RA Anand R.;
RT "Walking, cloning and mapping with YACs in 3q27. Localisation of 5
RT ESTs including 3 members of the Cystatin gene family and
RT identification of CpG islands.";
RL Genomics 32:425-430(1996).
DR EMBL: X88976; CAA61407.1; -;
FT NON_TER
SQ SEQUENCE 8 AA; 925 MW; FD5411A7376871E6 CRC64;

Query Match 81.8%; Score 9; DB 4; Length 8;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 5 LTR 7

RESULT 2

Q9T4Y2 PRELIMINARY; PRT; 8 AA.
ID Q9T4Y2
AC Q9T4Y2;
DT 01-MAY-2000 (TRENBLrel. 13, Created)

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE COI gene product (Fragment).
OS Aterina pectinifera (Starfish).
OC Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC Asterozoa; Valvatacea; Valvatida; Asterinidae; Asterina.
OX NCBI_TaxID=7594;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=89354669; PubMed=2766382;
RA Jacobs H.T., Asakawa S., Araki T., Miura K., Smith M.J., Watanabe K.;
RT "Conserved tRNA gene cluster in starfish mitochondrial DNA.";
RL Curr. Genet. 15:193-206(1989).
DR EMBL; X16886; CAA34767.1; -.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 8 AA; 1114 MW; F0C9D36415B736D6 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 8;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 3 LSR 5

RESULT 3
O89965
ID O89965 PRELIMINARY; PRT; 8 AA.
AC O89965;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Agnoprotein (Fragment).
OS Polyomavirus JC.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=10632;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=98244352; PubMed=9584961;
RA Boldorini R., Caldarelli-Stefano R., Monga G., Zocchi M., Mediat M.,
RA Tosoni A., Ferrante P.;
RT "PCR detection of JC virus DNA in the brain tissue of a 9-year-old
RT child with pleomorphic xanthoastrocytoma.";
RL J. Neurovirol. 4:242-245(1998).
DR EMBL; AF064547; AAC23995.1; -.
FT NON TER
SQ SEQUENCE 8 AA; 1002 MW; ED15B736C40732C6 CRC64;

Query Match 81.8%; Score 9; DB 12; Length 8;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 6 LSR 8

RESULT 4
Q97688
ID Q97688 PRELIMINARY; PRT; 9 AA.
AC Q97688;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Cytochrome c oxidase subunit I (fragment).
GN COI.
OS Gecko gecko (Tokay gecko).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae; Gekko.

```

```

OX NCBI_TaxID=36310;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=99343618; PubMed=10413626;
RA Macey J.R., Wang Y., Ananjeva N.B., Larson A., Papenfuss T.J.;
RT "Vicariant patterns of fragmentation among gekkonid lizards of the
RT genus teratoscincus produced by the indian collision: A molecular
RT phylogenetic perspective and an area cladogram for central asia.";
RL Mol. Phylogenet. Evol. 12:320-332(1999).
DR EMBL; AF114249; AAD51600.1; -.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 9 AA; 1188 MW; 428CB9C9D36411A7 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LTR 4

RESULT 5
Q94VH4
ID Q94VH4 PRELIMINARY; PRT; 9 AA.
AC Q94VH4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cytochrome c oxidase subunit I (fragment).
GN COI.
OS Varanus glauerti.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=169841;
[1]
RN SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407500; AAL10054.1; -.
KW Mitochondrion.
FT NON TER
SQ SEQUENCE 9 AA; 1124 MW; 9E80C733640DD731 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 3 LAR 5

RESULT 6
Q94VD8
ID Q94VD8 PRELIMINARY; PRT; 9 AA.
AC Q94VD8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cytochrome c oxidase subunit I (fragment).
GN COI.
OS Varanus niloticus.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=62046;
[1]
RN SEQUENCE FROM N.A.
RA Ast J.C.;

```

RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407514; AAL10096.1; -.
KW Mitochondrion.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1154 MW; 9E80C7336411A731 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 3 LTR 5

RESULT 7
Q94V18 ID Q94V18 PRELIMINARY; PRT; 9 AA.
AC Q94V18;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus eremias.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=169838;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407495; AAL10040.1; -.
KW Mitochondrion.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1124 MW; 9E80C733640DD731 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 3 LAR 5

RESULT 8
Q94VC6 ID Q94VC6 PRELIMINARY; PRT; 9 AA.
AC Q94VC6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus pilbarensis.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=62048;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407518; AAL10108.1; -.
KW Mitochondrion.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1064 MW; 874CASA36411A735 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 3 LTR 5

RESULT 9
Q94VE1 ID Q94VE1 PRELIMINARY; PRT; 9 AA.
AC Q94VE1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus mertensi.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=62044;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407512; AAL10090.1; -.
KW Mitochondrion.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1154 MW; 9E80C7336411A731 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 9;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 3 LTR 5

RESULT 10
Q91BM8 ID Q91BM8 PRELIMINARY; PRT; 9 AA.
AC Q91BM8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Papovavirus BK (Gardner derived clone BKV9) early transcription control region (Fragment).
OS Simian virus 12.
OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
OX NCBI_TaxID=46771;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Gardner;
RX MEDLINE=87061221; PubMed=3023684;
RA Chuke W.F., Walker D.L., Peitzman L.B., Frisque R.J.;
RT "Construction and characterization of hybrid polyomavirus genomes.";
RL J. Virol. 60:960-971(1986).
DR EMBL; M14452; AAA96236.1; -.
FT NON_TER 9
SQ SEQUENCE 9 AA; 1130 MW; C7FD15B736C40732 CRC64;

Query Match 81.8%; Score 9; DB 12; Length 9;
Best Local Similarity 66.7%; Pred. No. 8.3e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
DB 6 LSR 8

RESULT 11

Q9PYKI ID Q9PYKI PRELIMINARY; PRT; 9 AA.
 AC Q9PYKI; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Papovavirus BK (Gardner) early transcription control region (Fragment).
 OS Simian virus 12.
 OC Viruses; dsDNA viruses, no RNA stage; Polyomaviridae; Polyomavirus.
 OX NCBI_TaxID=46771;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Gardner;
 RA MEDLINE=87061221; PubMed=3023684;
 RA Chuks W.F., Walker D.L., Peitzman L.B., Frisque R.J.;
 RT "Construction and characterization of hybrid polyomavirus genomes.";
 RL J. Virol. 60:960-971(1986).
 DR EMBL; M14451; AAA96235.1; -.
 FT NON TER 9
 SQ SEQUENCE 9 AA; 1130 MW; C7FD15B736C40732 CRC64;

Query Match 81.8%; Score 9; DB 12; Length 9;
 Best Local Similarity 66.7%; Pred. No. 8.3e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
 |
 Db 6 LSR 8

RESULT 12

Q958K9 ID Q958K9 PRELIMINARY; PRT; 10 AA.
 AC Q958K9; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Rana boylei.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
 OX NCBI_TaxID=160499;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21184280; PubMed=1128498;
 RA Macey J.R., Strasburg J.L., Brisson J.A., Vredenburg V.T.,
 RA Jennings M., Larson A.;
 RT "Molecular Phylogenetics of Western North American Frogs of the Rana boylei Species Group."
 RL Mol. Phylogenet. Evol. 19:131-143(2001).
 DR EMBL; AF314019; AAK56877.1; -.
 KW Mitochondrion.
 FT NON TER 10
 SQ SEQUENCE 10 AA; 1320 MW; 42D380C9D36411A7 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 10;
 Best Local Similarity 66.7%; Pred. No. 3.8e+03;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
 |
 Db 3 LTR 5

RESULT 13

Q9TG86 ID Q9TG86 PRELIMINARY; PRT; 10 AA.
 AC Q9TG86;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Diploglossus bilobatus.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguillidae;
 OC Diploglossus.
 OX NCBI_TaxID=102183;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99343613; PubMed=10413621;
 RA Macey J.R., Schulte J.A. II, Larson A., Tunney B.S., Orlov N.,
 RA Papenfuss T.J.;
 RT "Molecular phylogenetics, tRNA evolution, and historical biogeography in anigulid lizards and related taxonomic families.";
 RL Mol. Phylogenet. Evol. 12:250-272(1999).
 DR EMBL; AF085608; AAD51514.1; -.
 KW Mitochondrion.
 FT NON TER 10
 SQ SEQUENCE 10 AA; 1255 MW; 5DEE80C7336411A7 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 10;
 Best Local Similarity 66.7%; Pred. No. 3.8e+03;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
 |
 Db 3 LTR 5

RESULT 14

Q94V97 ID Q94V97 PRELIMINARY; PRT; 10 AA.
 AC Q94V97; 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Cytochrome c oxidase subunit I (Fragment).
 GN COI.
 OS Varanus spenceri.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
 OX NCBI_TaxID=169854;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ast J.C.;
 RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
 RL Cladistics 17:0-0(2001).
 DR EMBL; AF407530; AAL10142.1; -.
 KW Mitochondrion.
 FT NON TER 10
 SQ SEQUENCE 10 AA; 1255 MW; 5DEE80C7336411A7 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 10;
 Best Local Similarity 66.7%; Pred. No. 3.8e+03;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
 |
 Db 3 LTR 5

RESULT 15

Q94VDS ID Q94VDS PRELIMINARY; PRT; 10 AA.
 AC Q94VDS;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Cytochrome c oxidase subunit I (Fragment).
GN COI.
OS Varanus olivaceus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidossauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=62047;
RN [1]
RP SEQUENCE FROM N.A.
RA Aet J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL Cladistics 17:0-0(2001).
DR EMBL; AF407515; AAL10099.1; -.
KW Mitochondrion.
FT NON TER 10 10
SQ SEQUENCE 10 AA; 1234 MW; 584C4CA5A36411A7 CRC64;

Query Match 81.8%; Score 9; DB 8; Length 10;
Best Local Similarity 66.7%; Pred. No. 3.8e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
Db 3 LTR 5

Search completed: December 8, 2003, 09:18:42
Job time : 6.20202 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 1.05051 Seconds
(without alignments)
179.063 Million cell updates/sec

Title: US-09-498-556C-357
Perfect score: 11
Sequence: 1 LXR 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	%	ID	Description
1	9	81.8	11	RS30_ONCMY	P83328 oncorhynchu
2	9	81.8	15	ONC1_ONCMY	P83287 oncorhynchu
3	9	81.8	15	UC25_MAIZE	P80631 zea mays (m
4	9	81.8	16	IBP4_PIG	P24854 sus scrofa
5	9	81.8	18	VAA5_RHOPA	Q02005 rhodopseudo
6	9	81.8	20	LYC_FELCA	P37155 felis silve
7	9	81.8	20	RECA_AZOVI	P37863 azotobacter
8	9	81.8	21	SYA_RAT	P50475 rattus norv
9	9	81.8	22	ANFC_CHICK	P21805 gallus gall
10	9	81.8	23	GLNA_PHOLP	P20479 phormidium
11	9	81.8	25	IPYR_PSEAN	P80898 pseudanabae
12	9	81.8	25	NEUL_RANTE	P20056 rana tempor
13	9	81.8	25	SPIG_PSEUS	P82357 pseudacanth
14	9	81.8	25	UBLI_BOVIN	P23356 bos taurus
15	9	81.8	26	NTRC_RHILP	P41502 rhizobium 1
16	9	81.8	26	YFHA_KLEPN	P21710 klebsiella
17	9	81.8	27	CKXK_CONRA	P58806 conus radia
18	9	81.8	27	SECR_CANFA	P09910 canis fami
19	9	81.8	27	SECR_RABIT	P32647 oryctolagus
20	9	81.8	27	SECR_SHEEP	P31299 ovis aries
21	9	81.8	28	V103_VACCP	Q00334 vaccinia vi
22	9	81.8	29	HS98_NEUCR	P31540 neurospora
23	9	81.8	29	Y51_EPT3	P20326 bacterioph
24	9	81.8	31	A98A_DROME	O46201 drosophila
25	9	81.8	31	DIUX_DIPPU	P82372 diptoptera
26	9	81.8	33	LYC2_HORSE	P81710 equus cabal
27	9	81.8	33	OTCC_PSEPU	P11727 pseudomonas
28	9	81.8	33	PK1_DICDI	P34101 dictyosteli
29	9	81.8	34	RNL1_PIG	P15466 sus scrofa
30	9	81.8	36	F4RE_METOG	P80951 methanogeni
31	9	81.8	36	PAHO_CERSI	P37999 ceratotheri
32	9	81.8	36	PAHO_CHIBR	P41519 chinchilla
33	9	81.8	36	PAHO_DIDMA	P18107 didelphis m

34 9 81.8 36 1 PAHO_EQUZE
35 9 81.8 36 1 PAHO_ERIEU
36 9 81.8 36 1 PAHO_MACMU
37 9 81.8 36 1 PAHO_RABIT
38 9 81.8 36 1 PAHO_TAPPI
39 9 81.8 36 1 PYU_AMICA
40 9 81.8 36 1 PYU_PIG
41 9 81.8 37 1 CALI_PIG
42 9 81.8 37 1 CALI_SHEEP
43 9 81.8 37 1 CALR_RANRI
44 9 81.8 37 1 PIF7_BOVIN
45 9 81.8 37 1 PSBY_GUITH

P38000 equus zebra
P41335 erinaceus e
P33884 macaca mula
P41336 oryctolagus
P39659 tapirus pin
P29205 amia calva
P01305 sus scrofa
P30880 sus scrofa
P30881 ovis aries
P31888 rana ridibu
P21671 bos taurus
O78433 guillardia

ALIGNMENTS

RESULT 1
RS30_ONCMY
ID RS30_ONCMY STANDARD; PRT; 11 AA.
AC P83328;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 40S ribosomal protein S30 (Fragment).
GN PAU.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]_TaxID=8022;
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE=Skin mucus;
RX MEDLINE=22142142; PubMed=12147245;
RA Fernandes J.M.O., Smith V.J.;
RT "A novel antimicrobial function for a ribosomal peptide from rainbow trout skin";
RL Biochem. Biophys. Res. Commun. 296:167-171(2002).
CC -!- FUNCTION: Has antibacterial activity against Gram-positive bacteria.
CC -!- MASS SPECTROMETRY: MW=6676.6; METHOD=MALDI.
CC -!- SIMILARITY: BELONGS TO THE S30E FAMILY OF RIBOSOMAL PROTEINS.
KW Ribosomal protein; Antibiotic.
FT NON_TER 11 11
SQ SEQUENCE 11 AA; 1123 MW; 2312AB630DD735B8 CRC64;
Query Match 81.8%; Score 9; DB 1; Length 11;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LXR 3
DB 6 LAR 8
RESULT 2
ONC1_ONCMY STANDARD; PRT; 15 AA.
ID ONC1_ONCMY
AC P83287;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC TISSUE=Skin;
RA Fernandes J.M.O., Smith V.J., Kemp G.D.;

```

RT "Purification and N-terminal sequencing of a 3 kDa antibacterial
RL peptide from skin secretions of rainbow trout."
CC Submitted (MAY-2002) to the SWISS-PROT data bank.
CC -I- FUNCTION: Has antibacterial activity against Gram-positive
CC bacterium P.citreus.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Skin.
DR GO: GO:000576; C:extracellular; NAS.
DR GO: GO:0003795; F:antimicrobial peptide activity; NAS.
DR GO: GO:0006805; P:xenobiotic metabolism; NAS.
KW Antibiotic.
FT UNSURE 4 4 OR G.
FT UNSURE 9 9 OR T.
FT NON_TER 15 15
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1601 MW; 43C25028DBC12B7C CRC64;

Query Match 81.8%; Score 9; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.4e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 12 LAR 14

RESULT 3
UC25_MAIZE STANDARD; PRT; 15 AA.
AC P80631;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 39, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 77)
DE (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernellet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 4.9. ITS MW IS: 31.6 kDa.
DR Maize-2DPAGE; P80631; COLEOPTILE.
DR MaizeDB; 123957; -.
FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1580 MW; 83C54CFOCE1614D0 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.4e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 5 LSR 7

RESULT 4
IBP4_PIG STANDARD; PRT; 16 AA.
AC P24854;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Insulin-like growth factor binding protein 4 (IGFBP-4) (IBP-4)
DE (IGF-binding protein 4) (Fragment).

```

```

GN IGFBP4.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=92109718; PubMed=1722398;
RA Coleman M.E., Fan Y.-C.E., Ehterton T.D.;
RT "Identification and NH2-terminal amino acid sequence of three
RT insulin-like growth factor-binding proteins in porcine serum.";
RL Biochem. Biophys. Res. Commun. 181:1131-1136(1991).
CC -I- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFs
CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH
CC PROMOTING EFFECTS OF THE IGFs ON CELL CULTURE. THEY ALTER THE
CC INTERACTION OF IGFs WITH THEIR CELL SURFACE RECEPTORS.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- SIMILARITY: Contains 1 IGFBP domain.
CC -I- SIMILARITY: Contains 1 thyroglobulin type-I domain.
DR PIR; JH0517; JH0517.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000716; Thyroglobulin_1.
DR PROSITE; PS00222; IGF BINDING; PARTIAL.
DR PROSITE; PS00484; THYROGLOBULIN_1; PARTIAL.
KW Growth factor binding
FT NON_TER 16 16
SQ SEQUENCE 16 AA; 1799 MW; 4098884009655E2 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 16;
Best Local Similarity 66.7%; Pred. No. 5.8e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 14 LAR 16

RESULT 5
YAA5_RHOPA STANDARD; PRT; 18 AA.
AC Q02005;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in aaDr 5'region (Fragment).
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009;
RX MEDLINE=92394882; PubMed=1522059;
RA Dispensa M., Thomas C.T., Kim M.K., Perrotta J.A., Gibson J.,
RA Harwood C.S.;
RT "Anaerobic growth of Rhodopseudomonas palustris on 4-hydroxybenzoate
RT is dependent on AaDr, a member of the cyclic AMP receptor protein
RT family of transcriptional regulators.";
RL J. Bacteriol. 174:5803-5813(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M92426; AAA26089.1; -.
DR PIR; A43334; A43334.
KW Hypothetical protein.
FT NON_TER 1 1
SQ SEQUENCE 18 AA; 2173 MW; 61DABDFE4E317B0A CRC64;

```

Query Match 81.8%; Score 9; DB 1; Length 18;
 Best Local Similarity 66.7%; Pred. No. 6.6e+02;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LXR 3
 DB 6 LTR 8

RESULT 6
 LYC_FELCA STANDARD; PRT; 20 AA.
 AC P37155;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lysozyme C (EC 3.2.1.17) (1,4-beta-N-acetylmuramidase C) (Fragment).
 GN Lyz.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Milk;
 RX MEDLINE=90263403; PubMed=2344734;
 RA Halliday J.A., Bell K., McKenzie H.A., Shaw D.C.;
 RT "Feline whey proteins: identification, isolation and initial
 RT characterization of alpha-lactalbumin, beta-lactoglobulin and
 RT lysozyme";
 RL Comp. Biochem. Physiol. 95B:773-779 (1990).
 CC -!- FUNCTION: Lysozymes have primarily a bacteriolytic function; those
 CC in tissues and body fluids are associated with the monocyte-
 CC macrophage system and enhance the activity of immunogens.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-
 CC acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan
 CC heteropolymers of the prokaryotes cell walls.
 CC -!- SUBUNIT: Monomer.
 CC -!- MISCELLANEOUS: Lysozyme C is capable of both hydrolysis and
 CC transglycosylation; it shows also a slight esterase activity. It
 CC acts rapidly on both peptide-substituted and unsubstituted
 CC peptidoglycan, and slowly on chitin oligosaccharides.
 CC -!- SIMILARITY: Belongs to family 22 of glycosyl hydrolases.
 DR PIR; A60525; A60525.
 DR HSP; P11376; 2EOL.
 DR InterPro; IPR001916; Glyco_hydro_22.
 DR Pfam; PF00062; Iys; 1.
 DR PROSITE; PS00128; LACTALBUMIN_LYSOZYME; PARTIAL.
 KW Hydrolase; Glycosidase; Bacteriolytic enzyme; Milk.
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2314 MW; EBB824EA59425E13 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 20;
 Best Local Similarity 66.7%; Pred. No. 7.3e+02;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LXR 3
 DB 8 LAR 10

RESULT 7
 RECX_AZOVI STANDARD; PRT; 20 AA.
 AC P37863;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Regulatory protein recX (Fragment).
 GN RECX.
 OS Azotobacter vinelandii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Azotobacter.
 OX NCBI_TaxID=354;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92225347; PubMed=1563632;
 RA Venkatesh T.V., Das H.K.;
 RT "The Azotobacter vinelandii recA gene: sequence analysis and
 RT regulation of expression";
 RL Gene 113:47-53 (1992).
 RN [2]
 RP IDENTIFICATION.
 RX MEDLINE=94218258; PubMed=8165147;
 RA de Mot R., Schoofs G., Vanderleyden J.;
 RT "A putative regulatory gene downstream of recA is conserved in gram-
 RT negative and gram-positive bacteria";
 RL Nucleic Acids Res. 22:1313-1314 (1994).
 CC -!- FUNCTION: Modulates recA activity (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- SIMILARITY: BELONGS TO THE RECX FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S96898; -; NOT ANNOTATED_CDS.
 DR HAMAP; MF_01114; -; 1.
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2111 MW; C809F8BCED6CB56 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 20;
 Best Local Similarity 66.7%; Pred. No. 7.3e+02;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LXR 3
 DB 18 LAR 20

RESULT 8
 SYA_RAT STANDARD; PRT; 21 AA.
 ID SYA_RAT
 AC P50475;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlaRS)
 DE (Fragment).
 GN AARS.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=91249799; PubMed=2040280;
 RA Dignam J.D., Dignam S.S., Brumley L.L.;
 RT "Alanyl-tRNA synthetase from Escherichia coli, Bombyx mori and Ratus
 RT ratus. Existence of common structural features";
 RL Eur. J. Biochem. 198:201-210 (1991).
 CC -!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA (Ala) = AMP +
 CC diphosphate + L-alanyl-tRNA (Ala).
 CC -!- SUBUNIT: Monomer.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 DR PIR; S16073; S16073.
 DR InterPro; IPR006193; tRNA synt Ala.
 DR PROSITE; PS00860; AA TRNA_LIGASE II ALA; PARTIAL.
 KW Aminoacyl-tRNA synthetase; Protein Biosynthesis; Ligase; ATP-binding.

```
FT NON TER 1 1
FT NON TER 21 21
SQ SEQUENCE 21 AA; 2293 MW; D739DDC62CD43375 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 21;
Best Local Similarity 66.7%; Pred. No. 7.7e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 18 LAR 20

RESULT 9
ID ANFC_CHICK STANDARD; PRT; 22 AA.
AC P21805;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-type natriuretic peptide (CNP).
GN NPFC.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX TISSUE=Brain;
RX MEDLINE=91113186; PubMed=1989595;
RA Arimura J.J., Minamino N., Kangawa K., Matsuo H.;
RT "Isolation and Identification of C-type natriuretic peptide in
RT chicken brain.";
RL Biochem. Biophys. Res. Commun. 174:142-148(1991).
CC -!- FUNCTION: VASORELAXANT ACTIVITY. HAS A CGMP-STIMULATING ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE NATRIURETIC PEPTIDES FAMILY.
DR PIR; JT0581; JT0581.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP; 1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
KW Vasoactive.
FT DISULFID 6 22
SQ SEQUENCE 22 AA; 2244 MW; FC2A4706D8DAC025 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 22;
Best Local Similarity 66.7%; Pred. No. 8.1e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 2 LSR 4

RESULT 10
ID GLNA_PHOLP STANDARD; PRT; 23 AA.
AC P20479;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase)
DE (Fragment).
OS Phormidium lapideum.
OC Bacteria; Cyanobacteria; Oscillatoriales; Phormidum.
OX NCBI_TaxID=32060;
RN [1]
RP SEQUENCE.
RX MEDLINE=89214011; PubMed=2907514;
RA Sawa Y., Ochiai H., Yoshida K., Tanizawa K., Tanaka H., Soda K.;

RT "Glutamine synthetase from a cyanobacterium, Phormidium lapideum:
RT purification, characterization, and comparison with other
RT cyanobacterial enzymes.";
RL J. Biochem. 104:917-923(1988).
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -!- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO
CC HEXAGONS.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
DR PIR; PX0011; PX0011.
DR InterPro; IPR001691; GLN_synth.
DR PROSITE; PS00180; GLNA_1; PARTIAL.
DR PROSITE; PS00181; GLNA_ATP; PARTIAL.
KW Ligase.
FT NON TER 23 23
SQ SEQUENCE 23 AA; 2656 MW; 20B69C164D2A5739 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 23;
Best Local Similarity 66.7%; Pred. No. 8.5e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 7 LSR 9

RESULT 11
ID IPYR_PSEAN STANDARD; PRT; 25 AA.
AC P80898;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
DE hydrolase) (pPase) (Fragment).
GN PPA.
OS Pseudanabaena sp. (strain PCC 6901).
OC Bacteria; Cyanobacteria; Oscillatoriales; Pseudanabaena.
OX NCBI_TaxID=47918;
RN [1]
RP SEQUENCE.
RA Gomez R., Serrano A.;
RL Submitted (FEB-1997) to the SWISS-PROT data bank.
CC -!- FUNCTION: Hydrolyzes PPI generated in anabolic reactions.
CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.
CC -!- COFACTOR: Requires the presence of divalent metal cation.
CC Magnesium confers the highest activity. Binds 4 divalent cations
CC per subunit (By similarity).
CC -!- SUBUNIT: Homohexamer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the pPase family.
DR HAMAP; MF_00209; -; 1.
KW Hydrolase; Magnesium.
FT NON TER 25 25
SQ SEQUENCE 25 AA; 2630 MW; FB6138F3AED4D43E CRC64;

Query Match 81.8%; Score 9; DB 1; Length 25;
Best Local Similarity 66.7%; Pred. No. 9.3e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 3 LSR 5

RESULT 12
ID NEUU_RANTE STANDARD; PRT; 25 AA.
AC P20056;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
```

DE Neuromedin U-25 (NmU-25).
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
RN NCBI_TaxID=8407;
[1]
RP
RC
RQ
SEQUENCE=Intestine;
RC
MEDLINE=90078173; PubMed=2592357;
RA Domin J., Yiangou Y.G., Spokes R.A., Aitken A., Parmar K.B.,
RA Chrysanthou B.J., Bloom S.R.;
RT "The distribution, purification, and pharmacological action of an
amphibian neuromedin U";
RL J. Biol. Chem. 264:20881-20885(1989).
CC -!- FUNCTION: STIMULATES UTERINE SMOOTH MUSCLE CONTRACTION AND CAUSES
SELECTIVE VASOCONSTRICTION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE NMU FAMILY.
DR PIR; A34179; A34179.
DR InterPro; IPR001942; NMU.
DR SMART; SM00084; NMU; 1.
DR PROSITE; PS00967; NMU; 1.
KW Amidation; Hormone.
FT MOD RES 25 25
SQ SEQUENCE 25 AA; 2832 MW; 6A01D89F6DA06FD4 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 25;
Best Local Similarity 66.7%; Pred. No. 9.3e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
|
|
|
DB 14 LSR 16

RESULT 13
SPIG_PSEUS
ID SPIG_PSEUS STANDARD; PRT; 25 AA.
AC P82357;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Spingerin.
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Isoptera; Termitidae;
OC Macrotermitinae; Pseudacanthotermes.
OX NCBI_TaxID=115113;
[1]
RN
RQ
SEQUENCE, MASS SPECTROMETRY, AND FUNCTION.
RC TISSUE=Blood, and Salivary gland;
RX PubMed=11053427;
RA Lamberty M., Zachary D., Lanot R., Bordereau C., Robert A.,
RA Hoffmann J.A., Bulet P.;
RT "Insect Immunity. Constitutive expression of a cysteine-rich
antifungal and a linear antibacterial peptide in a termite insect.";
RL J. Biol. Chem. 276:4085-4092(2001).
CC -!- FUNCTION: ACTIVE AGAINST GRAM-POSITIVE BACTERIA B.MEGATERIUM AND
M.LUTEUS, GRAM-NEGATIVE BACTERIA E.COLI SB363 AND D22,
K.PNEUMONIAE, S.TYPHIMURIUM AND P.AERUGINOSA, YEAST C.ALBICANS
AND FILAMENTOUS FUNGI F.CULMORUM, N.CRASSA, N.HEMATOCOCCA AND
T.VIRIDAE. INACTIVE AGAINST GRAM-POSITIVE BACTERIA B.SUBTILIS,
S.PYOGENES, B.THURINGIENSIS AND S.AUREUS, GRAM-NEGATIVE BACTERIA
E.CLOACAE AND E.CAROTOFORA AND FILAMENTOUS FUNGUS B.BASSIANA.
CC -!- INDUCTION: By bacterial infection.
CC -!- MASS SPECTROMETRY: MW=3001.8; METHOD=WALDI.
CC -!- MISCELLANEOUS: THERE ARE THREE ISOFORMS OF SPINGERIN.
KW Antibiotic; Fungicide.
FT VARIANT 1 3 MISSING (IN N-3 ISOFORM).
FT VARIANT 22 25 MISSING (IN C-4 ISOFORM).
SQ SEQUENCE 25 AA; 3001 MW; AA79370264262F60 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 25;
Best Local Similarity 66.7%; Pred. No. 9.3e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LXR 3
|
|
|
DB 22 LTR 24

RESULT 14
UBLI_BOVIN
ID UBLI_BOVIN STANDARD; PRT; 25 AA.
AC P23356;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase isozyme L1 (EC 3.4.19.12) (UCH-
L1) (Ubiquitin thiolesterase L1) (Neuron cytoplasmic protein 9.5)
DE UCHL1.
GN UCHL1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovioidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RQ
SEQUENCE
RC TISSUE=Brain;
RX MEDLINE=92008646; PubMed=1833240;
RA Giambanco I., Bianchi R., Ceccarelli P., Pula G., Sorci G.,
RA Antonoli S., Bocchini V., Donato R.;
RT "Neuron-specific" protein gene product 9.5 (PGP 9.5) is also
expressed in glioma cell lines and its expression depends on cellular
growth state.";
RL FEBS Lett. 290:131-134(1991).
CC -!- FUNCTION: UBIQUITIN-PROTEIN HYDROLASE IS INVOLVED BOTH IN THE
PROCESSING OF UBIQUITIN PRECURSORS AND OF UBIQUITINATED PROTEINS.
CC THIS ENZYME IS A THIOL PROTEASE THAT RECOGNIZE AND HYDROLYZE
A PEPTIDE BOND AT THE C-TERMINAL GLYCINE OF UBIQUITIN.
CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
ubiquitin + a thiol.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: NEURONS AND CELLS OF THE DIFFUSE
NEUROENDOCRINE SYSTEM AND THEIR TUMORS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C12.
DR PIR; S17561; S17561.
DR MEROPS; C12.001; -.
DR InterPro; IPR001578; UCH 1.
DR Pfam; PF01088; Peptidase C12; 1.
DR PROSITE; PS00140; UCH 1, PARTIAL.
KW Ubl conjugation pathway; Hydrolase; Thiol protease; Multigene family.
FT NON TER 25 25
SQ SEQUENCE 25 AA; 2812 MW; 26BB5ADD0A754D55 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 25;
Best Local Similarity 66.7%; Pred. No. 9.3e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LXR 3
|
|
|
DB 17 LTR 19

RESULT 15
NTRC_RHILP
ID NTRC_RHILP STANDARD; PRT; 26 AA.
AC P41502;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nitrogen assimilation regulatory protein (Fragment).
GN NTRC.
OS Rhizobium leguminosarum (biovar phaseoli).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=385;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CE-3;
RX MEDLINE=94018651; PubMed=8412703;
RA Patriarca E.J., Riccio A., Tate R., Colonna-Romano S., Iaccarino M.,
RA Defez R.;
RT "The ntrBC genes of Rhizobium leguminosarum are part of a complex
operon subject to negative regulation.";
RL Mol. Microbiol. 9:569-577(1993).
CC -!- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM NTRB/NTRC
CC INVOLVED IN THE ACTIVATION OF NITROGEN ASSIMILATORY GENES SUCH AS
CC GUNA. NTRC IS PHOSPHORYLATED BY NTRB AND INTERACTS WITH SIGMA-54.
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X71436; -; NOT_ANNOTATED_CDS.
DR PIR; S36203;
DR InterPro; IPR001789; Response reg.
DR PROSITE; PS0110; RESPONSE REGULATORY; 1.
KW Nitrogen fixation; Transcription regulation; Repressor; Activator;
KW DNA-binding; ATP-binding; Phosphorylation; Sensory transduction.
FT DOMAIN 1 >26 RESPONSE REGULATORY.
FT NON_TER 26
SQ SEQUENCE 26 AA; 2687 MW; E11ECCG3C8G304 CRC64;

Query Match 81.8%; Score 9; DB 1; Length 26;
Best Local Similarity 66.7%; Pred. No. 9.7e+02;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3
Db 22 LSR 24

Search completed: December 8, 2003, 09:14:15
Job time : 2.05051 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2003, 09:13:30 ; Search time 1.65657 Seconds
(without alignments)
232.212 Million cell updates/sec

Title: US-09-498-556C-357

Perfect score: 11

Sequence: 1 LXR 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	81.8	7	E48394	glycoprotein compo
2	9	81.8	7	B48394	major fat-globule
3	9	81.8	8	PC4131	hypothetical prote
4	9	81.8	8	T13818	cytochrome oxidase
5	9	81.8	10	T13838	cytochrome-c oxida
6	9	81.8	10	T14219	cytochrome-c oxida
7	9	81.8	11	PU0034	dextranucrase (EC
8	9	81.8	12	A42324	cytochrome P450c27
9	9	81.8	13	PT0293	19 heavy chain CRD
10	9	81.8	14	PA0111	protein QA100054 -
11	9	81.8	15	S36896	ribosomal protein
12	9	81.8	15	PA0075	fructose-bisphosph
13	9	81.8	15	PA0102	fructose-bisphosph
14	9	81.8	15	B32800	hypothetical prote
15	9	81.8	15	PH1455	phe leader peptide
16	9	81.8	15	AF0832	T-cell receptor al
17	9	81.8	16	C61414	chymotrypsin (EC 3
18	9	81.8	16	JH0517	insulin-like growt
19	9	81.8	16	B44036	collagen alpha 1(X
20	9	81.8	17	JP0046	ribosomal protein
21	9	81.8	17	A34835	ribosomal protein
22	9	81.8	17	I54269	vitamin D binding
23	9	81.8	17	JQ2310	hypothetical 2.1K
24	9	81.8	17	JQ2320	hypothetical 2.1K
25	9	81.8	17	B61414	chymotrypsin (EC 3
26	9	81.8	18	S29491	GTP-binding protei
27	9	81.8	18	A41877	LcrKc - Yersinia p
28	9	81.8	18	A43334	orf1 5' of aadR -
29	9	81.8	18	S33645	hypothetical prote

30	9	81.8	18	2	S10452	hypothetical prote
31	9	81.8	19	2	S02269	glycogen(starch) s
32	9	81.8	19	2	UC2060	homeobox 5 protein
33	9	81.8	19	2	A60326	cholecystokinin-58
34	9	81.8	20	2	A60525	lysozyme (EC 3.2.1
35	9	81.8	20	2	JP0050	ribosomal protein
36	9	81.8	20	2	JP0051	hypothetical prote
37	9	81.8	20	2	S33001	serine proteinase
38	9	81.8	20	2	PC2084	globin - polychaet
39	9	81.8	20	2	S19616	chymotrypsin (EC 3
40	9	81.8	20	2	A61414	J-kappa recombinat
41	9	81.8	20	2	A42267	probable trp opero
42	9	81.8	20	2	AC0269	alanine-trna ligas
43	9	81.8	21	2	S16073	natriuretic peptid
44	9	81.8	22	2	JT0581	insulin-like growt
45	9	81.8	22	2	F23734	

ALIGNMENTS

RESULT 1

E48394

glycoprotein component 16/major fat-globule membrane protein/MFG-E8 homolog - bovine (fr
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 07-Feb-1997

C:Accession: E48394

R:Mather, I.H.; Banghart, L.R.; Lane, W.S.

Biochem. Mol. Biol. Int. 29, 545-554, 1993

A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig

II-like sequences.

A:Reference number: A48394; MUID:93250576; PMID:8485470

A:Accession: E48394

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-7 <MAT>

A:Experimental source: milk

A>Note: sequence extracted from NCBI backbone (NCBIP:131450)

C:Keywords: glycoprotein

Query Match 81.8%; Score 9; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3

Db 5 LAR 7

RESULT 2

B48394

major fat-globule membrane protein GP 55 - guinea pig (fragment)

C:Species: Cavia porcellus (guinea pig)

C:Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995

C:Accession: B48394

R:Mather, I.H.; Banghart, L.R.; Lane, W.S.

Biochem. Mol. Biol. Int. 29, 545-554, 1993

A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-pig

II-like sequences.

A:Reference number: A48394; MUID:93250576; PMID:8485470

A:Accession: B48394

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-7 <MAT>

A:Experimental source: milk

A>Note: sequence extracted from NCBI backbone (NCBIP:131444)

Query Match 81.8%; Score 9; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3

Db 5 LAR 7

Db 5 LAR 7

RESULT 3
PC4131
hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)
C;Species: Pseudomonas aeruginosa
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: PC4131
R;Kawabaki, S.; Arai, H.; Igarashi, Y.; Kodama, T.
Gene 167, 87-91, 1995
A;Title: Sequencing and characterization of the downstream region of the genes encoding Y for biosynthesis of heme d1.
A;Reference number: JC4552; MUID:96144254; PMID:8566817
A;Accession: PC4131
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-8 <RAW>
A;Cross-references: DDBJ:D50473; NID:g1217594
A;Note: this ORF is not annotated in GenBank entry PSENIRC, release 113.0

Query Match 81.8%; Score 9; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
| |
Db 4 LSR 6

RESULT 4
Tl3818
Cytochrome oxidase subunit I - Atlantic hagfish mitochondrion (fragment)
C;Species: mitochondrion Myxine glutinosa (Atlantic hagfish)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: Tl3818
R;Delarbre, C.; Barriuel, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A;Title: The main features of the craniate mitochondrial DNA between the ND1 and the COI
A;Reference number: Z17775; MUID:97398704; PMID:9254918
A;Accession: Tl3818
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-8
A;Cross-references: EMBL:Y09527; NID:g2340019; PIDN:CRAA70718.1; PID:g2340022
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion

Query Match 81.8%; Score 9; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
| |
Db 3 LSR 5

RESULT 5
Tl3838
cytochrome-c oxidase (EC 1.9.3.1) chain I - Bipes biporus mitochondrion (fragment)
C;Species: mitochondrion Bipes biporus
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
C;Accession: Tl3838
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.
Mol. Biol. Evol. 14, 91-104, 1997
A;Title: Two novel gene orders and the role of light-strand replication in rearrangement
A;Reference number: Z17789; MUID:97153826; PMID:9000757
A;Accession: Tl3838
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>

A;Cross-references: EMBL:U71335; NID:g1753232; PID:g1753235; PIDN:AAB48271.1
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

Query Match 81.8%; Score 9; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
| |
Db 3 LTR 5

RESULT 6
Tl4219
cytochrome-c oxidase (EC 1.9.3.1) chain I - Xenosaurus grandis mitochondrion (fragment)
C;Species: mitochondrion Xenosaurus grandis
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: Tl4219
R;Macey, J.R.; Larson, A.; Ananjeva, N.B.; Fang, Z.; Papenfuss, T.J.
Mol. Biol. Evol. 14, 91-104, 1997
A;Title: Two novel gene orders and the role of light-strand replication in rearrangement
A;Reference number: Z17789; MUID:97153826; PMID:9000757
A;Accession: Tl4219
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-10 <MAC>
A;Cross-references: EMBL:U71333; NID:g5739536; PIDN:AAC62821.1; PID:g1753275
C;Genetics:
A;Genome: mitochondrion
A;Note: COI
C;Keywords: mitochondrion; oxidoreductase

Query Match 81.8%; Score 9; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
| |
Db 3 LTR 5

RESULT 7
PU0034
dextranucrase (EC 2.4.1.5) - Streptococcus bovis (fragment)
C;Species: Streptococcus bovis
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 18-Sep-1996
C;Accession: PU0034
R;Uezono, Y.; Tsumori, H.; Mukasa, H.
submitted to JIPID, October 1993
A;Description: Purification and properties of glucosyltransferase synthesizing 1,6-alpha-
A;Reference number: PU0034
A;Accession: PU0034
A;Molecule type: protein
A;Residues: 1-11 <UEZ>
A;Experimental source: ATCC 9809
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 81.8%; Score 9; DB 2; Length 11;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
| |
Db 8 LTR 10

RESULT 8
A42324
cytochrome P450c27/25 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)

C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Feb-1996
C;Accession: A42324
R;Shayiq, R.M.; Avadhani, N.G.
J. Biol. Chem. 267, 2421-2428, 1992
A;Title: Sequence complementarity between the 5'-terminal regions of mRNAs for rat mitochondria.
A;Reference number: A42324; MUID:92129322; PMID:1733943
A;Accession: A42324
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-12 <SHA>
A;Note: sequence extracted from NCBI backbone (NCBIN:78404, NCBIIP:78410)

Query Match 81.8%; Score 9; DB 2; Length 12;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
|
Db 4 LSR 6
|

RESULT 9
PT0293
Ig heavy chain CDR3 region (clone 4-124) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0293
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and J_H gene segments in the development of B lymphocytes.
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0293
A;Molecule type: DNA
A;Residues: 1-13 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotrimer; immunoglobulin

Query Match 81.8%; Score 9; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
|
Db 1 LSR 3
|

RESULT 10
PA0111
protein QAI00054 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 07-Apr-1995 #sequence_revision 26-May-1995 #text_change 06-Jun-1997
C;Accession: PA0111
R;Kamo, M.; Kawakami, T.; Tsugita, A.
submitted to JIPID, March 1995
A;Reference number: PA0109
A;Accession: PA0111
A;Molecule type: protein
A;Residues: 1-14 <KAM>
A;Experimental source: root

Query Match 81.8%; Score 9; DB 2; Length 14;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
|
Db 11 LSR 13
|

RESULT 11
S36896
ribosomal protein S16 - Mycobacterium bovis (fragment)

C;Species: Mycobacterium bovis
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Jan-1995
C;Accession: S36896
R;Ohara, N.; Kimura, M.; Higashi, Y.; Yamada, T.
FEBS Lett. 331, 9-14, 1993
A;Title: Isolation and amino acid sequence of the 30S ribosomal protein S19 from Mycobacterium bovis.
A;Reference number: S36887; MUID:94009653; PMID:8405418
A;Accession: S36896
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <OHA>
C;Keywords: protein biosynthesis; ribosome

Query Match 81.8%; Score 9; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
|
Db 6 LTR 8
|

RESULT 12
PA0075
fructose-bisphosphate aldolase (EC 4.1.2.13) I - fungus (Fusarium sporotrichioides) (fragment)
N;Alternate names: aldolase; fructose-1,6-bisphosphate triosephosphate-lase
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Mar-2001
C;Accession: PA0075; PA0077
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichioides.
A;Reference number: PA0051
A;Accession: PA0075
A;Molecule type: protein
A;Residues: 1-15 <CHO>
A;Note: this form (II) had a molecular weight of 30.6K and an isoelectric point of 5.4
C;Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 81.8%; Score 9; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
|
Db 6 LSR 8
|

RESULT 13
PA0102
fructose-bisphosphate aldolase (EC 4.1.2.13) III - fungus (Fusarium sporotrichioides) (fragment)
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Mar-2001
C;Accession: PA0102
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrichioides.
A;Reference number: PA0051
A;Accession: PA0102
A;Molecule type: protein
A;Residues: 1-15 <CHO>
C;Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 81.8%; Score 9; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXR 3
|
Db 6 LSR 8
|

Db 6 LSR 8

RESULT 14

B32800
hypothetical protein (P1 5' region) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 30-Sep-1993
C/Accession: B32800
R/Jindal, S.; Dudani, A.K.; Singh, B.; Harley, C.B.; Gupta, R.S.
Mol. Cell. Biol. 9, 2279-2283, 1989
A:Title: Primary structure of a human mitochondrial protein homologous to the bacterial
A:Reference number: A32800; MUID:89313783; PMID:2568584
A/Accession: B32800
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-15 <JIN>
A/Cross-references: GB:M22382

Query Match 81.8%; Score 9; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3

Db 3 LSR 5

RESULT 15

PH1455
T-cell receptor alpha chain (clone A24/PEF4) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995
C/Accession: PH1455
R/Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; K
J. Exp. Med. 177, 811-820, 1993
A:Title: T cell receptor selection by and recognition of two class I major histocompatib
A:Reference number: PH1430; MUID:93171821; PMID:8436911
A/Accession: PH1455
A:Molecule type: mRNA
A:Residues: 1-15 <CAS>
A/Experimental source: cytolytic T-lymphocyte
C:Superfamily: immunoglobulin homology
C:Keywords: receptor; T-cell

Query Match 81.8%; Score 9; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.7e+03;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LXR 3

Db 3 LSR 5

Search completed: December 8, 2003, 09:19:34
Job time : 2.65657 secs